



# STIC Search Report

## Biotech-Chem Library

STIC Database Tracking Number: 123567

**TO:** Celine Qian  
**Location:** REM-2C70/2A89  
**Art Unit:** 1636  
**Friday, June 04, 2004**

**Case Serial Number:** 09/515363

**From:** Edward Hart  
**Location:** Biotech-Chem Library  
REM-1A55  
**Phone:** 571-272-2512

**edward.hart@uspto.gov**

### Search Notes

Examiner Qian,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Edward Hart



# STIC SEARCH RESULT FEEDBACK FORM

## Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher or contact:*

Mary Hale, Information Branch Supervisor  
571-272-2507 Remsen E01 D86

### Voluntary Results Feedback Form

➤ *I am an examiner in Workgroup:*  *Example: 1610*

➤ *Relevant prior art found, search results used as follows:*

- 102 rejection
- 103 rejection
- Cited as being of interest.
- Helped examiner better understand the invention.
- Helped examiner better understand the state of the art in their technology.

*Types of relevant prior art found:*

- Foreign Patent(s)
- Non-Patent Literature  
(journal articles, conference proceedings, new product announcements etc.)

➤ *Relevant prior art not found:*

- Results verified the lack of relevant prior art (helped determine patentability).
- Results were not useful in determining patentability or understanding the invention

Comments:

Drop off or send completed forms to STIC/Biotech-Chem Library Remsen Bldg.



Pending Nucleic Acid and Pending Amino Acid database searches generate two sets of results each. The Pending Nucleic Acid and Pending Amino Acid database searches generate two sets of results each. The Pending databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches. Searches run against the Nucleic Acid Pending database produce two sets of results, with the extensions **.rnpm** and **.rnpn**. Searches run against the Amino Acid Pending database produce two sets of results, with the extensions **.rapm** and **.rapn**.

***Because they contain data that is confidential, the results of Pending database searches should not be left in the case .***

677N: Ed Hart.

123552

Access DB# \_\_\_\_\_

SEARCH REQUEST FORM

Scientific and Technical Information Center

CBFE

Requester's Full Name Celine Qian Examiner # 78710 Date: 6/2/04  
Art Unit 1636 Phone Number 572-0777 Serial Number 09/515363  
Mail Box and Bldg Room Location 2189 Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention Melanoma differentiation associated gene-5

Inventors (please provide full names): Paul Fisher et al.

Earliest Priority Filing Date 2/29/2000

\*For Sequence Searches Only\* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number

Please search for SEQ ID NO: 1 & 2.

STAFF USE ONLY	Type of Search	Vendors and cost where applicable
Searcher _____	NA Sequence (#) <u>1</u>	STN _____
Searcher Phone # _____	AA Sequence (#) <u>1</u>	Dialog _____
Searcher Location _____	Structure (#) _____	Qwestic, Other _____
Date Searcher Prepared <u>6/2/04</u>	Bibliographic _____	DBI, etc. _____
Date Entered <u>6/4/04</u>	Litigation _____	Lexis, Nexis _____
Searcher Prep & Review (#) _____	Full Text _____	Sequence Systems <u>04/05P</u>
Editorial Prep (#) _____	Patent Family _____	WWW Internet _____
Editorial Review (#) _____	Other _____	Other Spec'd _____

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

protein - protein search, using sw model

on: June 2, 2004, 19:00:27 ; Search time 63 Seconds  
(without alignments)  
4597.003 Million cell updates/sec

title: US-09-515-363C-2

perfect score: 5311

sequence: 1 MSGNGYSTDENRYLISCFRA.....LPIITFPNLDYSECCLFSDD 1025

scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

searched: 1586107 seqs, 282547505 residues

maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

AAE10155;  
29-NOV-2001 (first entry)  
Human melanoma differentiation associated (Mda)-5 protein.  
Human; melanoma differentiation associated gene; Mda-5; interferon; IFN;  
RNA helicase motif; caspase recruitment domain; CARD; therapy; melanoma;  
neuroblastoma; astrocytoma; glioblastoma; multiforme; cancer; cervical;  
breast; colon; prostate; osteosarcoma; chondrosarcoma; systemic toxicity;  
central nervous system; cytostatic; apoptosis.

red. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

卷之三

1	5311	100.0	1025	4	AAE10155	Human mel	FT	/note= "ATPase A motif"
2	5285	99.5	1025	5	AAM47798	Human RNA	FT	443 . . 446
3	5285	99.5	1025	7	ADC31794	Human nov	FT	/note= "ATPase B motif"
4	4306	81.1	838	4	AAE10165	RNA helic	FT	722 . . 823
5	2595.5	48.9	558	4	AAU23090	Novel hum	FT	/note= "Reveals significant homology to RNA helicase C-terminal conserved domain"
6	2456	46.2	468	6	ADA54899	Human pro	FT	
7	1860	35.0	417	4	AAU23647	Novel hum	FT	
8	1839	34.6	356	4	AAM40129	Novel hum	XX	WO200164707-A1.
9	1769	33.3	348	4	AAM41915	Human pol	XX	
10	1769	33.3	348	7	ADC33339	Human nov	PD	07-SEP-2001.
11	1352	25.5	678	5	ABP69672	Human pol	XX	
12	1344	25.3	678	4	AAB93708	Human pro	PF	28-FEB-2001; 2001WO-US006960.
13	1093	20.6	304	4	AAU00296	Interfero	XX	
14	1093	20.6	308	4	AAU00298	Interfero	PR	29-FEB-2000; 2000US-00515363.
15	861	16.2	267	4	AAU00297	Interfero	XX	
16	861	16.2	271	4	AAU00299	Interfero	PA	(UYCO ) UNIV COLUMBIA NEW YORK.
17	808	15.2	448	4	AAU23654	Novel hum	XX	
18	783	14.7	166	4	ABB42219	Peptide #	PI	Fisher PB, Kang D, Gopalkrishnan RV;
19	783	14.7	166	4	AAM36026	Peptide #	XX	
20	783	14.7	166	4	AAM75915	Human bon	DR	WPI; 2001-565494/63.
21	783	14.7	166	4	AAM63104	Human bra	DR	N-PSDB; AAD17203.
22	782	14.7	154	4	AAG74314	Human col	XX	
23	782	14.7	154	5	ABP41112	Human ova	PT	Nucleic acid sequences encoding a Melanoma Differentiation Associated
24	666.5	12.5	357	3	AAB42981	Human ORF	PT	Gene useful for cancer cell growth suppression, apoptosis and anti-viral
25	563.5	10.6	447	5	ABP41171	Human ORF	PT	activities.

XX	Claim 21; Page 18-19; 152pp; English.	QY	721 GIIFTKTROSAYALSOWTENKEFAEVGYKAHLIGAGHSSEFKPMTQNEQKEVISKFR	780
PS		Db	721 GIIFTKTROSAYALSOWTENKEFAEVGYKAHLIGAGHSSEFKPMTQNEQKEVISKFR	780
XX	The present invention relates to an isolated nucleic acid encoding a melanoma differentiation associated gene-5 (Mda-5) polypeptide. Mda-5 contains a caspase recruitment domain (CARD) and a RNA helicase motif. Mda-5 is a novel interferon (IFN) inducible gene with structural similarities to RNA helicases and CARD motif containing proteins. Mda-5 is induced during terminal differentiation in human melanoma cells treated with the combination of recombinant fibroblast IFN and the anti-leukaemic compound mezerein (MEZ). Mda-5 is useful for identifying compounds that may induce its expression. Mda-5 is useful for treating cancer e.g., melanoma, neuroblastoma, glioblastoma, astrocytoma, prostate, cervical cancer, breast cancer, colon cancer, prostate cancer, osteosarcoma, chondrosarcoma and especially where the cancer is a cancer of the central nervous system and apoptosis. The Mda-5 promoter exhibits melanocyte tissue specificity and minimises systemic toxicity. The present sequence is human Mda-5 protein.	QY	781 GRINLLIATTVAEGLDIKECNIVIRYGLVTNEIAMVQARGRAADESTYVLVAHSGSV	840
CC		Db	781 GRINLLIATTVAEGLDIKECNIVIRYGLVTNEIAMVQARGRAADESTYVLVAHSGSV	840
CC		QY	841 IEHETVNDFREKMYKAHVQNMKEEYAHKILELQLOMOSIMECKMKTKRNIAKHYKNNP	900
CC		Db	841 IEHETVNDFREKMYKAHVQNMKEEYAHKILELQLOMOSIMECKMKTKRNIAKHYKNNP	900
CC		QY	901 SLITFLCKNCNSVLAQSGEDIHVIERKHHVNMTPEFKELYIVRENKAQKCADYQINGEI	960
CC		Db	901 SLITFLCKNCNSVLAQSGEDIHVIERKHHVNMTPEFKELYIVRENKAQKCADYQINGEI	960
CC		QY	961 ICKCGQAWGTMVHKGLDPLCILKIRNFVYVFKNNSTKKQYKKWELPITFPNLDSSECL	1020
CC		Db	961 ICKCGQAWGTMVHKGLDPLCILKIRNFVYVFKNNSTKKQYKKWELPITFPNLDSSECL	1020
CC	Sequence 1025 AA;	QY	1021 FSDED 1025	
XX		Db	1021 FSDED 1025	
SQ	Query Match Score 5311; DB 4; Length 1025; Best Local Similarity 100.0%; Pred. No. 0; Matches 1025; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	RESULT 2 AAM47798		
Db	1 MSNGYSTDENFRLISCFRARVYOMYIQLVEPVLDYLTLFLPAEVKEQIORTVATSGNMQAVE 60	ID AAM47798		
QY	1 MSNGYSTDENFRLISCFRARVYOMYIQLVEPVLDYLTLFLPAEVKEQIORTVATSGNMQAVE 60	XX		
Db	61 ILLSTLERKGWHLGWTRFVEALRRTGSPSPLAARYMNPELTDLPSFSENAHDEYLQLLNL 120	AC AAM47798;		
QY	61 ILLSTLERKGWHLGWTRFVEALRRTGSPSPLAARYMNPELTDLPSFSENAHDEYLQLLNL 120	XX		
Db	121 LQPTLVDKLLVRDVLDKCMCMBEELLTIEDRNRRIAAEENNNGNEGVRELLKRVQKENWFA 180	DT 05-MAR-2002 (First entry)		
QY	121 LQPTLVDKLLVRDVLDKCMCMBEELLTIEDRNRRIAAEENNNGNEGVRELLKRVQKENWFA 180	XX		
Db	121 LQPTLVDKLLVRDVLDKCMCMBEELLTIEDRNRRIAAEENNNGNEGVRELLKRVQKENWFA 180	DE Human RNA helicase RH116.		
QY	181 FLNVLRQTGNNELVQELTGSDCCSESNAEIEINLSQLDGPQVEEQLLSTTVQPNLKEVWGM 240	KW Human; RH116; RNA helicase; cytostatic; virucide; anti-HIV; immunosuppressive; immunostimulatory; antiarthritic; antiarteriosclerotic; osteopathic; antidiabetic; hepatotropic; antiinflammatory; cancer; infection; HIV; hepatitis; genetic disease; autoimmune disease; graft rejection; vaccine.		
Db	181 FLNVLRQTGNNELVQELTGSDCCSESNAEIEINLSQLDGPQVEEQLLSTTVQPNLKEVWGM 240	XX		
QY	241 ENNSSESSFADSSVVSESDTSLAEGSVSCLDESLGHNNSNMGSDSGTMGSDSDDEENVAARA 300	KW Homo sapiens.		
Db	241 ENNSSESSFADSSVVSESDTSLAEGSVSCLDESLGHNNSNMGSDSGTMGSDSDDEENVAARA 300	XX		
QY	301 SPEPELQLRQYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDHLDKCKKQASEPGKVI 360	PD 15-NOV-2001.		
Db	301 SPEPELQLRQYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDHLDKCKKQASEPGKVI 360	XX		
QY	361 VLVNKVLLVEQLFRKEFQPFLKWRVIGLSGDTQLKISFPEVVKSCDIIISTAQILENS 420	PR 11-MAY-2001; 2000FR-00006030.		
Db	361 VLVNKVLLVEQLFRKEFQPFLKWRVIGLSGDTQLKISFPEVVKSCDIIISTAQILENS 420	XX		
QY	421 LLNLNGEDAGVQLSDFLSLLIIDECHHTNKEAVYNNIMRHMLMQLRNRLKENKPVIP 480	PA (ISTA-) INSTAC.		
Db	421 LLNLNGEDAGVQLSDFLSLLIIDECHHTNKEAVYNNIMRHMLMQLRNRLKENKPVIP 480	PA (INSP ) INST PASTEUR LILLE.		
QY	481 LPQIQLGILTASPGVGGATKOAKAAEEHILKLCANLDAFTIKTVKENLDQLKQIOPCKKFA 540	XX		
Db	481 LPQIQLGILTASPGVGGATKOAKAAEEHILKLCANLDAFTIKTVKENLDQLKQIOPCKKFA 540	PI Bahr G, Cocude C, Capron A;		
QY	541 IADATREDPFKEKLLEIMTRIQTQYQMSMSPMSDFGTQPEQWA1QMEKKAAKKGKRNKERV	XX		
Db	541 IADATREDPFKEKLLEIMTRIQTQYQMSMSPMSDFGTQPEQWA1QMEKKAAKKGKRNKERV	XX		
QY	601 AEHLRKYNEALQINDTIRMIAYTHLETFYNEEKDKKFAVIEDDDSEGGDDEYCDGDEDE 660	CC The present sequence is the protein sequence for human RH116. RH116 is a 116kDa protein and has homology to RNA helicases (DEXH box). RH116 and		
Db	601 AEHLRKYNEALQINDTIRMIAYTHLETFYNEEKDKKFAVIEDDDSEGGDDEYCDGDEDE 660	CC its coding sequence are useful for treating cancer; acute or chronic		
QY	661 DDLKKPLKLDETDRFLMLFFENNKMKLRAENPEYENEKLTKLRTNTIMEQYTRTEESAR 720	CC infections (especially by HIV or hepatitis B or C); inherited genetic		
Db	661 DDLKKPLKLDETDRFLMLFFENNKMKLRAENPEYENEKLTKLRTNTIMEQYTRTEESAR 720	CC diseases; (auto)immune diseases (particularly rheumatism, arthritis, CC arteriosclerosis, osteoporosis and diabetes, but many others listed) and CC to prevent graft rejection. RH116 and its coding sequence are also useful CC for inducing, or increasing, the immune response to a vaccine		

XX	SQ	Sequence 1025 AA;	Db	961 ICKCGQAWGTMVHKGLDLPCLKIRNFVVVFKNNSTKKQYKKWELPITFPNLDYSECCL 1020
	Query	Match 99.5%; Score 5285; DB 5; Length 1025;	QY	1021 FSDED 1025
	Best Local Similarity 99.5%; Pred. No. 0;	Matches 1020; Conservative 2; Mismatches 3; Indels 0; Gaps 0;	DB	1021 FSDED 1025
QY	1	MSNGYSTDENFRYLISCFRARVKMYIQVEPVLDYLTLFLPAEVKEQIQRSTVATSGNMQAVE 60		RESULT 3
Db	1	MSNGYSTDENFRYLISCFRARVKMYIQVEPVLDYLTLFLPAEVKEQIQRSTVATSGNMQAVE 60		ADC31794
QY	61	LILLSTLERGVWHLGWTREFVEALRRRTGSPLAARYMNPELTDLPSPSPENAHDLEYQLNNL 120		ADC31794 standard; protein; 1025 AA.
Db	61	LILLSTLERGVWHLGWTREFVEALRRRTGSPLAARYMNPELTDLPSPSPENAHDLEYQLNNL 120		XX
QY	121	LQPTLYDKLLVVDVLDKCMEEELLTIEDRNRRIAAEENNNGESGVRELLKRIVKENWFA 180		XX
Db	121	LQPTLYDKLLVVDVLDKCMEEELLTIEDRNRRIAAEENNNGESGVRELLKRIVKENWFA 180		Human novel polypeptide sequence, SEQ ID NO:1876.
QY	181	FLNVLRQTNELVQELTGSDCSESNAEILNLSQVDGPOVEQLLSTVQPNLEKEVWGM 240		DE
Db	181	FLNVLRQTNELVQELTGSDCSESNAEILNLSQVDGPOVEQLLSTVQPNLEKEVWGM 240		Human; diagnostic; drug screening; forensics; gene mapping; neurodegenerative diseases; Parkinson's disease; Alzheimer's disease; ulcers; osteoporosis; autoimmune disease; anaemia; platelet disorder; wound; burns; molecular weight marker; food supplement; antiparkinsonian; nootropic; neuroprotective; anticoagulant; thrombolytic; pulmonary; nootropics; antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic; gene therapy.
QY	241	ENNSSESSFADSSVVSSESDTSLAEGSVSCLDESGLHNSNMGSDSDEENVAARA 300		AC
Db	241	ENNSSESSFADSSVVSSESDTSLAEGSVSCLDESGLHNSNMGSDSDEENVAARA 300		XX
QY	301	SEPEPLQRLPRPYQMEVAQPALEGKNNIICLPTGSGKTRAVVIAKDHLDKKKASEPGKVI 360		18-DEC-2003 (first entry)
Db	301	SEPEPLQRLPRPYQMEVAQPALEGKNNIICLPTGSGKTRAVVIAKDHLDKKKASEPGKVI 360		XX
QY	361	VLVNKVLLVEQLFRKEFQPFLLKKWVIGLSDTOLKISFPEVVKSCDIIISTAAQILENS 420		XX
Db	361	VLVNKVLLVEQLFRKEFQPFLLKKWVIGLSDTOLKISFPEVVKSCDIIISTAAQILENS 420		XX
QY	421	LNNLLENGEDAGVQLSDFSLIIIDECHTNKEAVYNNIMRHYLMQKLNRLKKENKPVIP 480		XX
Db	421	LNNLLENGEDAGVQLSDFSLIIIDECHTNKEAVYNNIMRHYLMQKLNRLKKENKPVIP 480		HYSEQ INC.
QY	481	LPQILGLTASPVGGAQKAAEHLKLCAKLDAFTIKTKVENLDQKLNQIQEPCKKFA 540		XX
Db	481	LPQILGLTASPVGGAQKAAEHLKLCAKLDAFTIKTKVENLDQKLNQIQEPCKKFA 540		XX
QY	541	IADATREDPPFKEKLLEIMTRIQTYCQMSPMSPMSDFGTQPYQWAIQMEKKAAKKKRNKRVC 600		10-APR-2003.
Db	541	IADATREDPPFKEKLLEIMTRIQTYCQMSPMSPMSDFGTQPYQWAIQMEKKAAKKKRNKRVC 600		XX
QY	601	AEHLRKYNEALQINDTIRMIDAYTHLETFYNEEKDKKFAVIEDDSDEGGDDEYCDGDEDE 660		24-SEP-2002; 2002WO-US030474.
Db	601	AEHLRKYNEALQINDTIRMIDAYTHLETFYNEEKDKKFAVIEDDSDEGGDDEYCDGDEDE 660		XX
QY	661	DLKKPLKLDDETDRFLMTLFFENKMLKRLAENPEYENKELTKLRTNTIMEQYTRTEESAR 720		24-SEP-2001; 2001US-0324631P.
Db	661	DLKKPLKLDDETDRFLMTLFFENKMLKRLAENPEYENKELTKLRTNTIMEQYTRTEESAR 720		XX
QY	721	GIIFTKTRQSAVALQWITENBKAFAEVGVKAHHLIGAGHSSEFKPMTNEQKEVISKFR 780		(HYSE-) HYSEQ INC.
Db	721	GIIFTKTRQSAVALQWITENBKAFAEVGVKAHHLIGAGHSSEFKPMTNEQKEVISKFR 780		XX
QY	781	GKINLLIATTVAEEGLDIKECNIVIRGLVNEIAMQARGRARADESTYYVLAHSGSGV 840		Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
Db	781	GKINLLIATTVAEEGLDIKECNIVIRGLVNEIAMQARGRARADESTYYVLAHSGSGV 840		Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
QY	841	IEHETVNDFREKQMYKAIAHCYQNMKPEEYAHKILEQMQSTIMEKKMKTKRNTIAKHYKNNP 900		Haley-Vicente D, Drmanac RT;
Db	841	IERETVNDFREKQMYKAIAHCYQNMKPEEYAHKILEQMQSTIMEKKMKTKRNTIAKHYKNNP 900		WPI; 2003-371981/35.
QY	901	SLITFLCKNCVLAQGEDIHVIEKMHMHVNMTPEFKELYIVRENKAHQKCADYQINGEI 960		N-PSDB; ADC30823.
Db	901	SLITFLCKNCVLAQGEDIHVIEKMHMHVNMTPEFKELYIVRENKAHQKCADYQINGEI 960		XX
QY	961	ICKCGQAWGTMVHKGLDLPCLKIRNFVVVFKNSTKKQYKKWELPITFPNLDYSECCL 1020		New polynucleotide and polypeptide useful for diagnosing, preventing or treating conditions such as neurodegenerative diseases, anemias, platelet disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or cancer.
QY	1	MSNGYSTDENFRYLISCFRARVKMYIQVEPVLDYLTLFLPAEVKEQIQRSTVATSGNMQAVE 60		Claim 20; SEQ ID NO 1876; 1185PP; English.
Db	1	MSNGYSTDENFRYLISCFRARVKMYIQVEPVLDYLTLFLPAEVKEQIQRSTVATSGNMQAVE 60		XX
QY	61	LILLSTLERGVWHLGWTREFVEALRRRTGSPLAARYMNPELTDLPSPSPENAHDLEYQLNNL 120		PI
Db	61	LILLSTLERGVWHLGWTREFVEALRRRTGSPLAARYMNPELTDLPSPSPENAHDLEYQLNNL 120		PI
QY	121	LQPTLYDKLLVVDVLDKCMEEELLTIEDRNRRIAAEENNNGESGVRELLKRIVKENWFA 180		PI
Db	121	LQPTLYDKLLVVDVLDKCMEEELLTIEDRNRRIAAEENNNGESGVRELLKRIVKENWFA 180		PI
QY	181	FLNVLRQTNELVQELTGSDCSESNAEILNLSQVDGPOVEQLLSTVQPNLEKEVWGM 240		XX
Db	181	FLNVLRQTNELVQELTGSDCSESNAEILNLSQVDGPOVEQLLSTVQPNLEKEVWGM 240		XX
QY	241	ENNSSESSFADSSVVSSESDTSLAEGSVSCLDESGLHNSNMGSDSDEENVAARA 300		XX
Db	241	ENNSSESSFADSSVVSSESDTSLAEGSVSCLDESGLHNSNMGSDSDEENVAARA 300		XX
QY	301	SEPEPLQRLPRPYQMEVAQPALEGKNNIICLPTGSGKTRAVVIAKDHLDKKKASEPGKVI 360		XX
Db	301	SEPEPLQRLPRPYQMEVAQPALEGKNNIICLPTGSGKTRAVVIAKDHLDKKKASEPGKVI 360		XX
QY	361	VLVNKVLLVEQLFRKEFQPFLLKKWVIGLSDTOLKISFPEVVKSCDIIISTAAQILENS 420		XX
Db	361	VLVNKVLLVEQLFRKEFQPFLLKKWVIGLSDTOLKISFPEVVKSCDIIISTAAQILENS 420		XX
QY	421	LNNLLENGEDAGVQLSDFSLIIIDECHTNKEAVYNNIMRHYLMQKLNRLKKENKPVIP 480		XX
Db	421	LNNLLENGEDAGVQLSDFSLIIIDECHTNKEAVYNNIMRHYLMQKLNRLKKENKPVIP 480		XX
QY	481	LPQILGLTASPVGGAQKAAEHLKLCAKLDAFTIKTKVENLDQKLNQIQEPCKKFA 540		XX
Db	481	LPQILGLTASPVGGAQKAAEHLKLCAKLDAFTIKTKVENLDQKLNQIQEPCKKFA 540		XX
QY	541	IADATREDPPFKEKLLEIMTRIQTYCQMSPMSPMSDFGTQPYQWAIQMEKKAAKKKRNKRVC 600		XX
Db	541	IADATREDPPFKEKLLEIMTRIQTYCQMSPMSPMSDFGTQPYQWAIQMEKKAAKKKRNKRVC 600		XX
QY	601	AEHLRKYNEALQINDTIRMIDAYTHLETFYNEEKDKKFAVIEDDSDEGGDDEYCDGDEDE 660		XX
Db	601	AEHLRKYNEALQINDTIRMIDAYTHLETFYNEEKDKKFAVIEDDSDEGGDDEYCDGDEDE 660		XX
QY	661	DLKKPLKLDDETDRFLMTLFFENKMLKRLAENPEYENKELTKLRTNTIMEQYTRTEESAR 720		PS
Db	661	DLKKPLKLDDETDRFLMTLFFENKMLKRLAENPEYENKELTKLRTNTIMEQYTRTEESAR 720		PS
QY	721	GIIFTKTRQSAVALQWITENBKAFAEVGVKAHHLIGAGHSSEFKPMTNEQKEVISKFR 780		XX
Db	721	GIIFTKTRQSAVALQWITENBKAFAEVGVKAHHLIGAGHSSEFKPMTNEQKEVISKFR 780		XX
QY	781	GKINLLIATTVAEEGLDIKECNIVIRGLVNEIAMQARGRARADESTYYVLAHSGSGV 840		XX
Db	781	GKINLLIATTVAEEGLDIKECNIVIRGLVNEIAMQARGRARADESTYYVLAHSGSGV 840		XX
QY	841	IEHETVNDFREKQMYKAIAHCYQNMKPEEYAHKILEQMQSTIMEKKMKTKRNTIAKHYKNNP 900		XX
Db	841	IERETVNDFREKQMYKAIAHCYQNMKPEEYAHKILEQMQSTIMEKKMKTKRNTIAKHYKNNP 900		XX
QY	901	SLITFLCKNCVLAQGEDIHVIEKMHMHVNMTPEFKELYIVRENKAHQKCADYQINGEI 960		XX
Db	901	SLITFLCKNCVLAQGEDIHVIEKMHMHVNMTPEFKELYIVRENKAHQKCADYQINGEI 960		XX
QY	961	ICKCGQAWGTMVHKGLDLPCLKIRNFVVVFKNSTKKQYKKWELPITFPNLDYSECCL 1020		XX

CC also used for treating diseases such as Parkinson's disease, Alzheimer's disease and other neurodegenerative diseases, anaemia, platelet disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or cancer. The nucleic acids may also be used as hybridisation probes or primers, and in the recombinant production of a protein. The polypeptides are also useful in generating antibodies, as molecular weight markers, and as food supplements. The present sequence represents a specifically claimed human polypeptide sequence of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp://wipo.int/pub/published\\_pct\\_sequences](ftp://wipo.int/pub/published_pct_sequences).

Sequence 1025 AA;

Query Match 99.5%; Score 5285; DB 7; Length 1025;  
Best Local Similarity 99.5%; Pred. No. 0;  
Matches 1020; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY	1 MSNGYSTDENFRYLISCFRARVRYM1QVEPVLDYLTFLPAEVKEQ1QRTVATSGNMQAVE 60	RESULT 4	Qy 841 IEHETVNDREKMYKA1HCVQNMKPEEYAHKILEQMSIMEKMRTRNIAKHYKNNP 900
Db	1 MSNGYSTDENFRYLISCFRARVRYM1QVEPVLDYLTFLPAEVKEQ1QRTVATSGNMQAVE 60	AAE10165	Db 841 IERETVNDREKMYKA1HCVQNMKPEEYAHKILEQMSIMEKMRTRNIAKHYKNNP 900
QY	61 LLLSTLEKGWHLGWTRFVEALRRTGSPLAARMNPELTDLPSFSFENAHDDEYLQLNNL 120	ID AAE10165 standard; protein; 838 AA.	Qy 901 SLITFLCKNCVSLACSGEDIHVIEKMHVNMTPEFKELYIVRENKAQKCADYQINGEI 960
Db	61 LLLSTLEKGWHLGWTRFVEALRRTGSPLAARYM1QVEPVLDYLTFLPAEVKEQ1QRTVATSGNMQAVE 60	XX	Db 901 SLITFLCKNCVSLACSGEDIHVIEKMHVNMTPEFKELYIVRENKAQKCADYQINGEI 960
QY	121 LQPTLVDKLLVRDVLDLKCMEEELLTIEDNRNRIAANNGNESGVRELLKRTIVOKENWFSA 180	AC AAE10165;	Qy 961 ICKGQAWGTMWVKGLDLPCLKIRNFVVVFKNINSTKCKQYKKWELPITFPNLDYSECC 1020
Db	121 LQPTLVDKLLVRDVLDLKCMEEELLTIEDNRNRIAANNGNESGVRELLKRTIVOKENWFSA 180	XX	Db 961 ICKGQAWGTMWVKGLDLPCLKIRNFVVVFKNINSTKCKQYKKWELPITFPNLDYSECC 1020
QY	181 FLNVLRQTGNNELVQELTGSDCSESNAE1ENLSQLDGPQVEEQLLSTTQPNLEKEVWGM 240	DT 29-NOV-2001 (first entry)	Qy 1021 FSDED 1025
Db	181 FLNVLRQTGNNELVQELTGSDCSESNAE1ENLSQLDGPQVEEQLLSTTQPNLEKEVWGM 240	XX	Db 1021 FSDED 1025
QY	241 ENNSSESSPADSSVVVSESDTSLAEGSVSCLIDESLGHNNSNMGSDSGTMGSDSDEENVAARA 300	DE RNA helicase conserved motif of human Mda-5 protein.	Qy 1021 FSDED 1025
Db	241 ENNSSESSPADSSVVVSESDTSLAEGSVSCLIDESLGHNNSNMGSDSGTMGSDSDEENVAARA 300	XX	Db 1021 FSDED 1025
QY	301 SPEPELQLRPFYQMEVAQPALEGKNI1ICLFTGSGKTRVAVYIAKDHLDKKKKKASEPGKVI 360	XX	Qy 1021 FSDED 1025
Db	301 SPEPELQLRPFYQMEVAQPALEGKNI1ICLFTGSGKTRVAVYIAKDHLDKKKKKASEPGKVI 360	XX	Db 1021 FSDED 1025
QY	361 VLVNKVLLVEQLFRKEFQPFLKWKWYRIGLISGDTQLKISFPEVVKSCDIIISTSTAQILENS 420	PF 28-FEB-2001; 2001WO-US006960.	Qy 1021 FSDED 1025
Db	361 VLVNKVLLVEQLFRKEFQPFLKWKWYRIGLISGDTQLKISFPEVVKSCDIIISTSTAQILENS 420	XX	Db 1021 FSDED 1025
QY	421 LLNLLENGEDAGVQLSDFSLIIIDECHHTNKEAVYNNIMRHYLMQKLKNRNLKVENKPVIP 480	PR 29-FEB-2000; 2000US-00515363.	Qy 1021 FSDED 1025
Db	421 LLNLLENGEDAGVQLSDFSFIIIDECHHTNKEAVYNNIMRHYLMQKLKNRNLKVENKPVIP 480	XX	Db 1021 FSDED 1025
QY	481 LPQILGLTASPGVGGATKQAKAEEHILKLCANLDAFTIKTYKENLQDQLKNO1QEPCKKFA 540	PA (UYCO ) UNIV COLUMBIA NEW YORK.	Qy 1021 FSDED 1025
Db	481 LPQILGLTASPGVGGATKQAKAEEHILKLCANLDAFTIKTYKENLQDQLKNO1QEPCKKFA 540	XX	Db 1021 FSDED 1025
QY	541 IADATREDPFKEKLIEMTRIQTYCQMSPMDSDEGTQPEQWA1QMEKKAQKGNRKERV 600	PI Fisher PB, Kang D, Gopalkrishnan RV;	Qy 1021 FSDED 1025
Db	541 IADATREDPFKEKLIEMTRIQTYCQMSPMDSDEGTQPEQWA1QMEKKAQKGNRKESVC 600	XX	Db 1021 FSDED 1025
QY	601 AEHLRKYNEALQINDTIRMIDAYTHLETFYNEEKDKKFAVIEDDSDEGGDDEYCDGDEDE 660	PT Gene useful for cancer cell growth suppression, apoptosis and anti-viral activity.	Qy 1021 FSDED 1025
Db	601 AEHLRKYNEALQINDTIRMIDAYTHLETFYNEEKDKKFAVIEDDSDEGGDDEYCDGDEDE 660	XX	Db 1021 FSDED 1025
QY	661 DDLKKPLKLDDETDRMLTFFENNMKMLKRIIAENPEYENELKLTKLRTNTIMEQYTRTEESAR 720	PS Example 1; Fig 1D; 152pp; English.	Qy 1021 FSDED 1025
Db	661 DDLKKPLKLDDETDRMLTFFENNMKMLKRIIAENPEYENELKLTKLRTNTIMEQYTRTEESAR 720	XX	Db 1021 FSDED 1025
QY	721 GIIFTKTRQSAVALSOWITENEKFAEVGVKAHLIGAHSSEFKEMTONEQKEVISKFR 780	CC The present invention relates to an isolated nucleic acid encoding a melanoma differentiation associated gene-5 (Mda-5) polypeptide. Mda-5 contains a caspase recruitment domain (CARD) and a RNA helicase motif.	Qy 1021 FSDED 1025
Db	721 GIIFTKTRQSAVALSOWITENEKFAEVGVKAHLIGAHSSEFKEMTONEQKEVISKFR 780	CC Mda-5 is a novel interferon (IFN) inducible gene with structural similarities to RNA helicases and CARD motif containing proteins. Mda-5 is induced during terminal differentiation in human melanoma cells treated with the combination of recombinant fibroblast IFN and the antileukaemic compound mezerein (MEZ). Mda-5 is useful for identifying compounds that may induce its expression. Mda-5 is useful for treating cancer e.g., melanoma, neuroblastoma, astrocytoma, glioblastoma multiforme, cervical cancer, breast cancer, colon cancer, prostate	Db 1021 FSDED 1025
QY	781 GKNLLIATTVAEGLDIKECNIVYGLVTNEIAMVQARGRAADESTYVLVAHSGSGV 840	CC cancer, osteosarcoma, chondrosarcoma and especially where the cancer is a cancer of the central nervous system and apoptosis. The Mda-5 promoter exhibits melanocyte tissue specificity and minimises systemic toxicity.	Qy 1021 FSDED 1025
Db	781 GKNLLIATTVAEGLDIKECNIVYGLVTNEIAMVQARGRAADESTYVLVAHSGSGV 840	CC The present sequence is RNA helicase conserved motif of human Mda-5 protein	Db 1021 FSDED 1025



PR 14-SEP-2000; 2000US-02333064P.  
 PR 14-SEP-2000; 2000US-02333065P.  
 PR 21-SEP-2000; 2000US-0234223P.  
 PR 21-SEP-2000; 2000US-0234274P.  
 PR 25-SEP-2000; 2000US-0234997P.  
 PR 25-SEP-2000; 2000US-0234998P.  
 PR 26-SEP-2000; 2000US-0235484P.  
 PR 27-SEP-2000; 2000US-0235834P.  
 PR 27-SEP-2000; 2000US-0235836P.  
 PR 29-SEP-2000; 2000US-0236327P.  
 PR 29-SEP-2000; 2000US-0236367P.  
 PR 29-SEP-2000; 2000US-0236368P.  
 PR 29-SEP-2000; 2000US-0236369P.  
 PR 29-SEP-2000; 2000US-0236370P.  
 PR 02-OCT-2000; 2000US-0236802P.  
 PR 02-OCT-2000; 2000US-0237037P.  
 PR 02-OCT-2000; 2000US-0237038P.  
 PR 02-OCT-2000; 2000US-0237039P.  
 PR 02-OCT-2000; 2000US-0237040P.  
 PR 13-OCT-2000; 2000US-0239935P.  
 PR 13-OCT-2000; 2000US-0239937P.  
 PR 20-OCT-2000; 2000US-0240960P.  
 PR 20-OCT-2000; 2000US-0241221P.  
 PR 20-OCT-2000; 2000US-0241785P.  
 PR 20-OCT-2000; 2000US-0241786P.  
 PR 20-OCT-2000; 2000US-0241787P.  
 PR 20-OCT-2000; 2000US-0241808P.  
 PR 20-OCT-2000; 2000US-0241809P.  
 PR 20-OCT-2000; 2000US-0241826P.  
 PR 01-NOV-2000; 2000US-0244617P.  
 PR 08-NOV-2000; 2000US-0246474P.  
 PR 08-NOV-2000; 2000US-0246475P.  
 PR 08-NOV-2000; 2000US-0246476P.  
 PR 08-NOV-2000; 2000US-0246477P.  
 PR 08-NOV-2000; 2000US-0246478P.  
 PR 08-NOV-2000; 2000US-0246523P.  
 PR 08-NOV-2000; 2000US-0246524P.  
 PR 08-NOV-2000; 2000US-0246525P.  
 PR 08-NOV-2000; 2000US-0246526P.  
 PR 08-NOV-2000; 2000US-0246527P.  
 PR 08-NOV-2000; 2000US-0246528P.  
 PR 08-NOV-2000; 2000US-0246532P.  
 PR 08-NOV-2000; 2000US-0246610P.  
 PR 08-NOV-2000; 2000US-0246611P.  
 PR 08-NOV-2000; 2000US-0246613P.  
 PR 17-NOV-2000; 2000US-0249207P.  
 PR 17-NOV-2000; 2000US-0249208P.  
 PR 17-NOV-2000; 2000US-0249209P.  
 PR 17-NOV-2000; 2000US-0249210P.  
 PR 17-NOV-2000; 2000US-0249211P.  
 PR 17-NOV-2000; 2000US-0249212P.  
 PR 17-NOV-2000; 2000US-0249213P.  
 PR 17-NOV-2000; 2000US-0249214P.  
 PR 17-NOV-2000; 2000US-0249215P.  
 PR 17-NOV-2000; 2000US-0249218P.  
 PR 17-NOV-2000; 2000US-0249244P.  
 PR 17-NOV-2000; 2000US-0249245P.  
 PR 17-NOV-2000; 2000US-0249264P.  
 PR 17-NOV-2000; 2000US-0249265P.  
 PR 17-NOV-2000; 2000US-0249297P.  
 PR 17-NOV-2000; 2000US-0249299P.  
 PR 01-DEC-2000; 2000US-0250160P.  
 PR 01-DEC-2000; 2000US-0250391P.  
 PR 05-DEC-2000; 2000US-0251030P.  
 PR 05-DEC-2000; 2000US-0251988P.  
 PR 06-DEC-2000; 2000US-0256719P.  
 PR 08-DEC-2000; 2000US-0251479P.  
 PR 08-DEC-2000; 2000US-0251868P.

PR 08-DEC-2000; 2000US-0251869P.  
 PR 08-DEC-2000; 2000US-0251989P.  
 PR 08-DEC-2000; 2000US-0251990P.  
 PR 11-DEC-2000; 2000US-0254097P.  
 PR 05-JAN-2001; 2000US-0259678P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Barash SC, Ruben SM;  
 XX  
 DR WPI; 2001-465566/50.  
 DR N-PSDB; AAS40960.  
 XX  
 PT Novel polypeptides and polynucleotides useful for diagnosing, preventing, treating neural, immune system, muscular, reproductive, pulmonary, cardiovascular, renal, proliferative disorders and cancerous diseases.  
 XX  
 PS Claim 11; SEQ ID NO 1086; 1180PP; English.  
 XX  
 CC The present invention relates to the isolation of novel human enzyme polypeptides, and the cDNA (AAS40785-AAS41684) and genomic sequences encoding them. The enzyme polypeptides of the invention may comprise the functional classes of oxidoreductases, transferases, hydrolases, lyases, isomerases or ligases. The sequences of the invention are useful in the diagnosis, treatment, prevention and/or prognosis of a wide range of disorders, including hyperproliferative disorders (e.g. cancer), immunodeficiency disorders (e.g. AIDS) autoimmune disorders (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis), blood-related disorders (e.g. haemophilia), reproductive disorders (e.g. infertility) and infectious disorders (e.g. Influenza). The polynucleotides of the invention can also be used in gene therapy. AAU22915-AAU23814 represent the novel human enzyme polypeptides of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 558 AA;  
 Query Match 48.9%; Score 2595.5; DB 4;  
 Best Local Similarity 94.6%; Pred. No. 5.7e-192;  
 Matches 510; Conservative 10; Mismatches 16; Indels 3; Gaps 2;  
 QY 294 ENVAARASPEPELQLRPPQMEVAQPALEGKNNIIICLPTGSKTRAVYIYAKDHLDDKKKA 353  
 1 ENVAARASPEPELQLRPPQMEVAQPALEGKNNIIICLPTGSKTRAVYIYAKDHLDDKKKA 60  
 QY 354 SEPGRKVTVLYNKVLLVEQLFRKEFQPFKLKQVLSDFSLIIIDC8HTNKEAVYNNIMRHYLMQQLKKNRNLKK 413  
 61 SEPGRKVTVLYNKVLLVEQLFRKEFQPFKLKQVLSDFSLIIIDC8HTNKEAVYNNIMRHYLMQQLKKNRNLKK 120  
 QY 414 AQILENSLLNLENGEDAVQQLSDFSLIIIDC8HTNKEAVYNNIMRHYLMQQLKKNRNLKK 473  
 121 AQILENSLLNLENGEDAVQQLSDFSLIIIDC8HTNKEAVYNNIMRHYLMQQLKKNRNLKK 180  
 QY 474 ENKPVIPILPQILGLTASPVGGGATQAKAEEHILKLCANLDAFTIKTVKENLDQLKQIQ 533  
 181 ENKPVIPILPQILGLTASPVGGGATQAKAEEHILKLCANLDAFTIKTVKENLDQLKQIQ 240  
 QY 534 EPCKKFAIADATREDPFKEKLIIMTRIOTYQMSPMSDFGTQPYEWAQIOMEKKAQKG 593  
 241 EPCKKFAIADATREDPFKEKLIIMTRIOTYQMSPMSDFGTQPYEWAQIOMEKKAQKG 300  
 QY 654 CDGDEDDELLKKPLKLDETDRLMLTFFENNNKMLKRLAENPEYENENKLTKLRNTIMEQYT 713  
 301 NRKERVCAEHILRKYNEALQINDTIRMLTFFENNNKMLKRLAENPEYENENKLTKLRNTIMEQYT 713  
 DB 361 CDGDEDDELLKKPLKLDETDRLMLTFFENNNKMLKRLAENPEYENENKLTKLRNTIMEQYT 420  
 QY 714 RTEESARGIIFTKTRQSAVALSQWITENEKFAEVGVKAHLIGAGHSSEFKPMTQNEQKE 773

Db	421	RTESARGIIFTKTRCSAYALSQWITENEKFAEVGKKAHLLIGAHSSEFKPMTQNQKE	480	QY	738	ITENEKFAEVGKKAHLLIGAHSSEFKPMTQNQKEVTSKERTGKINLLIATTVAEGLD	797
QY	774	VISKFRTGKINLLIATVAEGLIKEKNCNTVTRYGLVNEIAM-VQARGRARADESTYY	831	Db	181	ITENEKFAEVGKKAHLLIGAHSSEFKPMTQNQKEVTSKERTGRINLLIATTVAEGLD	240
Db	481	VISKFRTGKINLLIAPQWQKRWILK--NVTLXSVWVSSPMKYHGFARGRARADESTYY	537	QY	798	IKECNIVTRYGLVNEIAMVQARGRARADESTYYVLAHSGSGVIEHETVNDFREKMYKA	857
				Db	241	IKECNIVTRYGLVNEIAMVQARGRARADESTYYVLAHSGSGVIERETVNDFREKMYKA	300
RESULT 6				QY	858	THCVQNMKPEEYAHKILELQMQSIMKMTKRNIAKHYKNPNSLITFLCKNCVLAACSG	917
	ADA54899	Standard; protein; 468 AA.		Db	301	THCVQNMKPEEYAHKILELQMQSIMKMTKRNIAKHYKNPNSLITFLCKNCVLAACSG	360
AC	ADA54899;			QY	918	EDIHVIEKOMHHVNMTPEFKELYIVRENKAQKKCADYQINGEELICKCGQAWGTMMVHKGL	977
XX	XX	20-NOV-2003 (First entry)		Db	361	EDIHVIEKOMHHVNMTPEFKELYIVRENKLQKKCADYQINGEELICKCGQAWGTMMVHKGL	420
DE	Human protein,	SEQ ID 2467.		QY	978	DLPCLKIRNFVUVFKNNSTKKQYKKWELPITPPNLDYSECLFSDED	1025
XX	KW	Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic; Gene Therapy; human; secretory protein; membrane proteins; cancer; inflammatory disease; osteoporosis; neurological disease.		Db	421	DLPCLKIRNFVUVFKNNSTKKQYKKWELPITPPNLDYSECLFSDED	468
XX	OS	Homo sapiens.		RESULT 7			
XX	PN	EP1293569-A2.			AAU23647		
XX	PD	19-MAR-2003.			ID	AAU23647	standard; protein; 417 AA.
XX	PF	21-MAR-2002; 2002EP-00006586.			XX	AAU23647;	
XX	PR	14-SEP-2001; 2001JP-00328381.			XX	AC	
XX	PR	24-JAN-2002; 2002US-0350435P.			XX	DT	18-DEC-2001 (First entry)
XX	PA	(HELI-) HELIX RES INST.			XX	DE	Novel human enzyme polypeptide #733.
PA	PA	(REAS-) RES ASSOC BIOTECHNOLOGY.			XX		Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase; ligase; hyperproliferative disorder; immunodeficiency disorder; autoimmune disorder; neurological disorder; metabolic disorder; inflammatory disorder; cardiovascular disorder; reproductive disorder; blood-related disorder; infectious disorder; cytostatic; anti arthritic; nephrotropic; anticoagulant.
XX	PI	Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;			XX		Homo sapiens.
PI	Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;				XX		OS
PI	Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;				XX		XX
XX	WPI	2003-395539/38.			PN	W0200155301-A2.	
XX	DR	N-PSDB; ADA53260.			XX	XX	
XX	PT	New polynucleotides encoding full-length polypeptides, e.g. secretory and/or membrane proteins, useful for developing medicines for diseases in which the gene is involved, or as target molecules for gene therapy.			XX	XX	02-AUG-2001.
PT	PT	PT			XX	XX	02-AUG-2001.
PT	PT	PT			XX	XX	17-JAN-2001; 2001WO-US001239.
PS	PS	Claim 14; SEQ ID NO 2467; 205pp; English.			XX	XX	17-JAN-2001; 2001WO-US001239.
XX	XX	XX			PR	31-JAN-2000; 2000US-0179065P.	
CC	CC	CC			PR	04-FEB-2000; 2000US-0180628P.	
CC	CC	CC			PR	24-FEB-2000; 2000US-0184664P.	
CC	CC	CC			PR	02-MAR-2000; 2000US-0186350P.	
CC	CC	CC			PR	16-MAR-2000; 2000US-0189874P.	
CC	CC	CC			PR	17-MAR-2000; 2000US-0190076P.	
CC	CC	CC			PR	18-APR-2000; 2000US-0198123P.	
CC	CC	CC			PR	19-MAY-2000; 2000US-0205515P.	
XX	XX	XX			PR	07-JUN-2000; 2000US-0209467P.	
XX	XX	XX			PR	28-JUN-2000; 2000US-0214886P.	
XX	XX	XX			PR	30-JUN-2000; 2000US-0215135P.	
XX	XX	XX			PR	07-JUL-2000; 2000US-0216647P.	
XX	XX	XX			PR	07-JUL-2000; 2000US-0216880P.	
XX	XX	XX			PR	11-JUL-2000; 2000US-0217487P.	
XX	XX	XX			PR	11-JUL-2000; 2000US-0217496P.	
XX	XX	XX			PR	14-JUL-2000; 2000US-0220963P.	
XX	XX	XX			PR	26-JUL-2000; 2000US-0220964P.	
XX	XX	XX			PR	14-AUG-2000; 2000US-0224518P.	
XX	XX	XX			PR	14-AUG-2000; 2000US-0224519P.	
XX	XX	XX			PR	14-AUG-2000; 2000US-0225213P.	
XX	XX	XX			PR	14-AUG-2000; 2000US-0225214P.	
XX	XX	XX			PR	14-AUG-2000; 2000US-0225266P.	
XX	XX	XX			PR	14-AUG-2000; 2000US-0225267P.	
XX	XX	XX			PR	14-AUG-2000; 2000US-0225268P.	
XX	XX	XX			PR	14-AUG-2000; 2000US-0225270P.	
XX	XX	XX			PR	14-AUG-2000; 2000US-0225447P.	
RESULT 6	ADA54899			Query Match	46.2*	Score 2456; DB 6; Length 468;	
	ID	ADA54899		Best Local Similarity	99.1%	Pred. No. 2.8e-181;	
	XX	XX		Matches	464;	Conservative 2; Mismatches 2; Indels 0; Gaps 0;	
QY	558	MTRIQTQCMSPMSDEFGTOPYEQWAIQMEKKAAKKGNRKERVCAEHLKYNEALQINDTI	617				
Db	1	MTRIQTQCMSPMSDEFGTOPYEQWAIQMEKKAAKKGNRKERVCAEHLKYNEALQINDTI	60				
QY	618	RMDIDAYTHLETFYNEEKKFAVIEDDSDEGGDDEYCDGDEDDEKKPLKLDETDRLM	677				
Db	61	RMDIDAYTHLETFYNEEKKFAVIEDDSDEGGDDEYCDGDEDDEKKPLKLDETDRLM	120				
QY	678	TLPFENNKLKRLAENPEYENELKTLRNLTMEQYTRTEESARGLFTKTRQSYALSQW	737				
Db	121	TLPFENNKLKRLAENPEYENELKTLRNLTMEQYTRTEESARGLFTKTRQSYALSQW	180				

PR 14-AUG-2000; 2000US-0225757P.  
 PR 14-AUG-2000; 2000US-0225758P.  
 PR 14-AUG-2000; 2000US-0225759P.  
 PR 18-AUG-2000; 2000US-0226279P.  
 PR 22-AUG-2000; 2000US-0226681P.  
 PR 22-AUG-2000; 2000US-0226688P.  
 PR 22-AUG-2000; 2000US-0227182P.  
 PR 23-AUG-2000; 2000US-0227009P.  
 PR 30-AUG-2000; 2000US-0228924P.  
 PR 01-SEP-2000; 2000US-0229287P.  
 PR 01-SEP-2000; 2000US-0229343P.  
 PR 01-SEP-2000; 2000US-0229344P.  
 PR 01-SEP-2000; 2000US-0229345P.  
 PR 05-SEP-2000; 2000US-0229509P.  
 PR 05-SEP-2000; 2000US-0229513P.  
 PR 06-SEP-2000; 2000US-0230437P.  
 PR 06-SEP-2000; 2000US-0230438P.  
 PR 08-SEP-2000; 2000US-0231242P.  
 PR 08-SEP-2000; 2000US-0231243P.  
 PR 08-SEP-2000; 2000US-0231244P.  
 PR 08-SEP-2000; 2000US-0231413P.  
 PR 08-SEP-2000; 2000US-0231414P.  
 PR 08-SEP-2000; 2000US-0232080P.  
 PR 08-SEP-2000; 2000US-0232081P.  
 PR 12-SEP-2000; 2000US-02331968P.  
 PR 14-SEP-2000; 2000US-02332397P.  
 PR 14-SEP-2000; 2000US-02331414P.  
 PR 08-SEP-2000; 2000US-02332398P.  
 PR 14-SEP-2000; 2000US-02332399P.  
 PR 14-SEP-2000; 2000US-02332400P.  
 PR 14-SEP-2000; 2000US-02332401P.  
 PR 14-SEP-2000; 2000US-02333063P.  
 PR 14-SEP-2000; 2000US-02333064P.  
 PR 14-SEP-2000; 2000US-02333065P.  
 PR 21-SEP-2000; 2000US-0234223P.  
 PR 21-SEP-2000; 2000US-0234274P.  
 PR 25-SEP-2000; 2000US-0234997P.  
 PR 25-SEP-2000; 2000US-0234998P.  
 PR 26-SEP-2000; 2000US-0235484P.  
 PR 27-SEP-2000; 2000US-0235836P.  
 PR 27-SEP-2000; 2000US-0235837P.  
 PR 29-SEP-2000; 2000US-0236327P.  
 PR 29-SEP-2000; 2000US-0236367P.  
 PR 29-SEP-2000; 2000US-0236368P.  
 PR 29-SEP-2000; 2000US-0236369P.  
 PR 02-OCT-2000; 2000US-0236370P.  
 PR 02-OCT-2000; 2000US-0236802P.  
 PR 02-OCT-2000; 2000US-0237037P.  
 PR 02-OCT-2000; 2000US-0237038P.  
 PR 02-OCT-2000; 2000US-0237039P.  
 PR 02-OCT-2000; 2000US-0237040P.  
 PR 13-OCT-2000; 2000US-023935P.  
 PR 13-OCT-2000; 2000US-023937P.  
 PR 20-OCT-2000; 2000US-0240960P.  
 PR 20-OCT-2000; 2000US-0241221P.  
 PR 20-OCT-2000; 2000US-0241785P.  
 PR 20-OCT-2000; 2000US-0241786P.  
 PR 08-NOV-2000; 2000US-0241787P.  
 PR 08-NOV-2000; 2000US-0241808P.  
 PR 20-OCT-2000; 2000US-0241809P.  
 PR 20-OCT-2000; 2000US-0241826P.  
 PR 01-NOV-2000; 2000US-0244617P.  
 PR 08-NOV-2000; 2000US-024474P.  
 PR 08-NOV-2000; 2000US-024475P.  
 PR 08-NOV-2000; 2000US-0246476P.  
 PR 08-NOV-2000; 2000US-0246477P.  
 PR 08-NOV-2000; 2000US-0246478P.  
 PR 08-NOV-2000; 2000US-0246523P.  
 PR 08-NOV-2000; 2000US-0246524P.  
 PR 08-NOV-2000; 2000US-0246525P.  
 PR 08-NOV-2000; 2000US-0246527P.  
 PR 08-NOV-2000; 2000US-0246528P.  
 PR 08-NOV-2000; 2000US-0246532P.

PR 08-NOV-2000; 2000US-0246609P.  
 PR 08-NOV-2000; 2000US-0246610P.  
 PR 08-NOV-2000; 2000US-0246611P.  
 PR 08-NOV-2000; 2000US-0246613P.  
 PR 17-NOV-2000; 2000US-0249207P.  
 PR 17-NOV-2000; 2000US-0249208P.  
 PR 17-NOV-2000; 2000US-0249209P.  
 PR 17-NOV-2000; 2000US-0249210P.  
 PR 17-NOV-2000; 2000US-0249211P.  
 PR 17-NOV-2000; 2000US-0249212P.  
 PR 17-NOV-2000; 2000US-0249213P.  
 PR 17-NOV-2000; 2000US-0249214P.  
 PR 17-NOV-2000; 2000US-0249215P.  
 PR 17-NOV-2000; 2000US-0249216P.  
 PR 17-NOV-2000; 2000US-0249217P.  
 PR 17-NOV-2000; 2000US-0249218P.  
 PR 17-NOV-2000; 2000US-0249244P.  
 PR 17-NOV-2000; 2000US-0249245P.  
 PR 17-NOV-2000; 2000US-0249300P.  
 PR 17-NOV-2000; 2000US-0249301P.  
 PR 01-DEC-2000; 2000US-0249265P.  
 PR 05-DEC-2000; 2000US-0249297P.  
 PR 05-DEC-2000; 2000US-0249299P.  
 PR 05-DEC-2000; 2000US-0249300P.  
 PR 05-DEC-2000; 2000US-0250160P.  
 PR 05-DEC-2000; 2000US-0250391P.  
 PR 08-DEC-2000; 2000US-0251030P.  
 PR 08-DEC-2000; 2000US-0251198P.  
 PR 08-DEC-2000; 2000US-0256719P.  
 PR 08-DEC-2000; 2000US-0251479P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251989P.  
 PR 08-DEC-2000; 2000US-0251990P.  
 PR 11-DEC-2000; 2000US-0254097P.  
 PR 05-JAN-2001; 2001US-0259678P.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX PI Rosen CA, Barash SC, Ruben SM;  
 XX PI WPI; 2001-465566/50.  
 XX DR N-PSDB; AAS41517.  
 XX PT Novel polypeptides and polynucleotides useful for diagnosing, preventing, treating neural, immune system, muscular, reproductive, pulmonary, cardiovascular, renal, proliferative disorders and cancerous diseases.  
 XX PS Claim 11; SEQ ID NO 1643; 1180pp; English.  
 XX PA  
 XX PI  
 XX PI  
 XX DR  
 XX DR  
 XX PT  
 XX PT  
 XX PS  
 XX CC  
 CC The present invention relates to the isolation of novel human enzyme polypeptides, and the cDNA (AAS40785-AAS41684) and genomic sequences encoding them. The enzyme polypeptides of the invention may comprise the functional classes of oxidoreductases, transferases, hydrolases, lyases, isomerases or ligases. The sequences of the invention are useful in the diagnosis, treatment, prevention and/or prognosis of a wide range of disorders including hyperproliferative disorders (e.g. cancer), immunodeficiency disorders (e.g. AIDS) autoimmune disorders (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis), blood-related disorders (e.g. haemophilia), reproductive disorders (e.g. infertility) and infectious disorders (e.g. Influenza). The polynucleotides of the invention can also be used in gene therapy. AAU22915-AAU23814 represent the novel human enzyme polypeptides of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX SQ Sequence 417 AA;  
 XX SQ Query Match 35.0%; Score 1860; DB 4;  
 Best Local Similarity 90.8%; Pred. No. 3.6e-135;  
 XX SQ

Matches	365;	Conservative	13;	Mismatches	20;	Indels	4;	Gaps	3;	DR	N-PSDB; AA159285.
QY	442	I	D	E	C	H	T	N	K	E	A
Db	8	D	E	C	H	T	N	K	E	A	V
QY	502	A	E	E	H	I	L	C	A	N	L
Db	68	A	E	E	H	I	L	C	A	N	L
QY	561	Q	T	Y	C	Q	M	S	P	M	S
Db	127	Q	T	Y	C	Q	M	S	P	M	S
QY	562	Q	T	Y	C	Q	M	S	P	M	S
Db	128	Q	T	Y	C	Q	M	S	P	M	S
QY	621	A	Y	Y	L	E	T	F	N	K	P
Db	187	A	Y	Y	L	E	T	F	N	K	P
QY	681	A	Y	Y	L	E	T	F	N	K	P
Db	247	A	Y	Y	L	E	T	F	N	K	P
QY	682	E	N	N	K	M	L	K	R	A	E
Db	248	E	N	N	K	M	L	K	R	A	E
QY	741	E	K	F	A	E	V	G	V	K	H
Db	307	E	K	F	A	E	V	G	V	K	H
QY	801	E	K	F	A	E	V	G	V	K	H
Db	308	E	K	F	A	E	V	G	V	K	H
QY	842	N	I	V	R	G	L	V	T	N	E
Db	406	N	I	V	R	G	L	V	T	N	E
RESULT 8											
ID	AAM40129	standard;	protein;	356	AA.						
XX											
AC	AAM40129;										
XX											
DT	22-OCT-2001	(first entry)									
XX											
DE	Human polypeptide	SEQ ID NO 3274.									
XX											
KW	Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemokinetic; chemotactic; haemostatic; leukaemia.										
XX											
OS	Homo sapiens.										
XX											
PN	WO200153312-A1.										
XX											
PD	26-JUL-2001.										
XX											
PF	26-DEC-2000;	200000-0034263.									
XX											
PR	23-DEC-1999;	99US-00471275.									
PR	21-JAN-2000;	20000US-00488725.									
PR	25-APR-2000;	20000US-0052317.									
PR	20-JUN-2000;	20000US-00598042.									
PR	19-JUL-2000;	20000US-00620312.									
PR	03-AUG-2000;	20000US-00653450.									
PR	14-SEP-2000;	20000US-00662191.									
PR	19-OCT-2000;	20000US-00693036.									
PR	29-NOV-2000;	20000US-00727344.									
XX											
PA	(HYSEQ INC.										
XX											
PI	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;										
PI	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;										
PI	Zhou P, Goodrich R, Drmanac RT;										
XX											
DR	WPI; 2001-442253/47.										

Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries.

Example 5; SEQ ID NO 3274; 10078pp; English.

XX

The invention relates to human nucleic acids (AA157798-AA161369) and the encoded polypeptides (AAM38642-AAM42213) with nootropic, immunosuppressant and cytostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and localised neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: Immune system suppression, Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, assays for receptor activity, arthritis and inflammation, leukaemias and C.N.S disorders. Note: The sequence data for this patent did not form part of the printed specification

XX

Sequence 356 AA;

Query Match 34.6%; Score 1839; DB 4; Length 356;

Best Local Similarity 99.4%; Pred. No. 1-2e-133;

Matches 350; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 674 RFLMLTFENNKKMLKRLAENPEYENELKTLRNTNTIMEQYTRTEESARGLIFTKTROSAYA 733

DB 5 RFLMLTFENNKKMLKRLAENPEYENELKTLRNTNTIMEQYTRTEESARGLIFTKTROSAYA 64

QY 734 LSQWITENKEFAEVGVKAHHLIGAGHSSEFKPMTQNEQEVISKFRGKINLLIATVVAE 793

DB 65 LSQWITENKEFAEVGVKAHHLIGAGHSSEFKPMTQNEQEVISKFRGKINLLIATVVAE 124

QY 794 EGLDIKECNIVIRYGLVTNEIAMQARGRARADESTYVLAHSGSGVIEHETVNDREK 853

DB 125 EGLDIKECNIVIRYGLVTNEIAMQARGRARADESTYVLAHSGSGVIEHETVNDREK 184

QY 854 MYKAITHCYQNMKPEEYAHKILELOMOSIMEKKMKTKRNIAKHYKNNPSSLITFLCKNC 913

DB 185 MYKAITHCYQNMKPEEYAHKILELOMOSIMEKKMKTKRNIAKHYKNNPSSLITFLCKNC 244

QY 914 ACSGEDIHVIEKMHVNVMTPEFKELYVRENKALQKKCADYQINGEIIICKGQAWGTMV 973

DB 245 ACSGEDIHVIEKMHVNVMTPEFKELYVRENKTLQKKCADYQINGEIIICKGQAWGTMV 304

RESULT 9

AAM41915 standard; protein; 356 AA.

XX

ID AAM41915

AC AAM41915;

XX

AC AAM41915;

XX

DT 22-OCT-2001 (first entry)

XX

DB 305 HKGLDLPCLKIRNFVVFKNNSTKKQYKKWVELBTTFPNLDDYSECCFLPSDED 356

XX

DE Human polypeptide SEQ ID NO 6846.

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemokinetic; chemotactic; haemostatic; leukaemia.

KW Human polypeptide SEQ ID NO 6846.

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemokinetic; chemotactic; haemostatic; leukaemia.

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemokinetic; chemotactic; haemostatic; leukaemia.

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemokinetic; chemotactic; haemostatic; leukaemia.

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemokinetic; chemotactic; haemostatic; leukaemia.

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemokinetic; chemotactic; haemostatic; leukaemia.

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemokinetic; chemotactic; haemostatic; leukaemia.

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemokinetic; chemotactic; haemostatic; leukaemia.

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemokinetic; chemotactic; haemostatic; leukaemia.

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemokinetic; chemotactic; haemostatic; leukaemia.

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemokinetic; chemotactic; haemostatic; leukaemia.

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemokinetic; chemotactic; haemostatic; leukaemia.

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemokinetic; chemotactic; haemostatic; leukaemia.

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemokinetic; chemotactic; haemostatic; leukaemia.

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemokinetic; chemotactic; haemostatic; leukaemia.

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemokinetic; chemotactic; haemostatic; leukaemia.

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemokinetic; chemotactic; haemostatic; leukaemia.

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemokinetic; chemotactic; haemostatic; leukaemia.

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemokinetic; chemotactic; haemostatic; leukaemia.

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemokinetic; chemotactic; haemostatic; leukaemia.

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemokinetic; chemotactic; haemostatic; leukaemia.

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemokinetic; chemotactic; haemostatic; leukaemia.

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemokinetic; chemotactic; haemostatic; leukaemia.

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemokinetic; chemotactic; haemostatic; leukaemia.

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemokinetic; chemotactic; haemostatic; leukaemia.

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemokinetic; chemotactic; haemostatic; leukaemia.

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemokinetic; chemotactic; haemostatic; leukaemia.

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemokinetic; chemotactic; haemostatic; leukaemia.

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemokinetic; chemotactic; haemostatic; leukaemia.

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemokinetic; chemotactic; haemostatic; leukaemia.

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system

PN	WO200153312-A1.	Db	303 PCLKIRNFVVFKNNSTKKQYKKWVELPITRPNLIDYSECCFLFSDED 348
XX	26-JUL-2001.		
XX	PF 26-DEC-2000; 2000WO-US034263.		
XX	PR 23-DEC-1999; 99US-00471275.		
PR 21-JAN-2000; 2000US-00488725.			
PR 25-APR-2000; 2000US-00552317.			
PR 20-JUN-2000; 2000US-00598042.			
PR 19-JUL-2000; 2000US-00620312.			
PR 03-AUG-2000; 2000US-00653450.			
PR 14-SEP-2000; 2000US-00662191.			
PR 19-OCT-2000; 2000US-00693036.			
PR 29-NOV-2000; 2000US-00727344.			
XX	(HYSE-) HYSEQ INC.		
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D, Wang J, Yang Y, Zhang J, Zhao QA;			
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Zhou P, Goodrich R, Drmanac RT;			
PI DR WPI; 2001-442253/47; DR N-PSDB; AAI61071.			
XX	PT Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries.		
XX	PS Example 2; SEQ ID NO 6846; 10078pp; English.		
CC	The invention relates to human nucleic acids (AAI57798-AAI61369) and the encoded polypeptides (AAM38642-AAM42213) with nootropic, immunosuppressant and cytostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and localised neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: Immune system suppression, Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, assays for receptor activity, arthritis and inflammation, leukaemias and C.N.S disorders. Note: The sequence data for this patent did not form part of the printed specification.		
XX	Sequence 348 AA;		
Qy	Query Match 33.3%; Score 1769; DB 4; Length 348;		
Db	Best Local Similarity 97.4%; Pred. No. 3.1e-128; Matches 337; Conservative 2; Mismatches 7; Indels 0; Gaps 0;		
Qy	680 FFENNKMLKRLAENPEYENEKLTKLNRNTIMEQYTRTEESARGIIIFTKTRQSAVYALSQWIT 739		
Db	3 FLKTIKKLNRLAEPKYENEKLTKLNRNTIMEQYTRTEESARGIIIFTKTRQSAVYALSQWIT 62		
Qy	740 ENEKFAEVGVKAHHLIGAHSSEFKPMTQNEQKEVISKFRTGKINLLIATTVAEGLDIK 799		
Db	63 ENEKFAEVGVKAHHLIGAHSSEFKPMTQNEQKEVISKFRTGKINLLIATTVAEGLDIK 122		
Qy	800 ECNVIRYGLVTNELAMQARGRAADESTYVLVAHSGSGVIBHETVNDFREKMYKAH 859		
Db	123 ECNVIRYGLVTNELAMQARGRAADESTYVLVAHSGSGVIBHETVNDFREKMYKAH 182		
Qy	860 CVQNMKPEYAHKILELQMQSIMEKKMKTAKRNIAKHYKNNPSLITFLCKNCVLAQSGED 919		
Db	183 CVQNMKPEYAHKILELQMQSIMEKKMKTAKRNIAKHYKNNPSLITFLCKNCVLAQSGED 242		
Qy	920 IHVIEKMHVNMTPPEFKELYIVRENKALQKCADYQINGEFLICKCGOAWGTMVHKGLDL 979		
Db	243 IHVIEKMHVNMTPPEFKELYIVRENKTLQKCADYQINGEFLICKCGOAWGTMVHKGLDL 302		
Qy	980 PCLKIRNFVVFKNNSTKKQYKKWVELPITRPNLIDYSECCFLFSDED 1025		





XX OS Homo sapiens.  
 XX WO200118208-A2.  
 XX PD 15-MAR-2001.  
 XX PF 08-SEP-2000; 2000WO-US024704.  
 XX PR 08-SEP-1999; 99US-0152921P.  
 XX PR 20-OCT-1999; 99US-0160575P.  
 XX PR 20-JAN-2000; 2000US-0177104P.  
 XX PR 07-SEP-2000; 2000US-00656633.  
 XX PA (CURA-) CURAGEN CORP.  
 XX PA (BIOJ ) BIOGEN INC.  
 XX PI Peyman JA, Da Silva A, Hochman P, Hsu A;  
 XX WPI; 2001-235201/24.  
 DR N-PSDB; AAS01149.  
 XX PT New interferon induced polypeptides and polynucleotides, useful for the  
 PT diagnosis, prevention and treatment of immunological, cell proliferative  
 PT disorders, such as lupus erythematosus, cancer, stroke and Alzheimer's  
 disease.  
 XX PS Claim 1; Page 29-32; 134pp; English.  
 XX The sequence represents interferon induced polypeptide, IFN4. IFN nucleic  
 CC acids and polypeptides are useful for treating or preventing a pathology  
 CC associated with IFN polypeptide in a human. They are useful for  
 CC determining the presence of or predisposition to a disease associated  
 CC with altered levels of IFN polypeptide or polynucleotide.  
 CC acids, polypeptides and antibodies are useful for diagnosis, prevention  
 CC or treatment of variety of immunological and cell proliferative  
 CC disorders, such as autoimmune diseases e.g. lupus erythematosus,  
 CC immunodeficiency diseases such as acquired immunodeficiency syndrome  
 CC (AIDS), graft rejection, viral infections including hepatitis and human  
 CC immunodeficiency virus (HIV), immune-mediated glomerulonephritis,  
 CC haematologic diseases such as aplastic anaemia and chronic neutropenia  
 CC and cancer. In addition they are also useful for treating or diagnosing  
 CC various disorders associated with cell death, including myocardial  
 CC infarction, stroke, neurological diseases including Alzheimer's and  
 CC Parkinson's diseases, amyotrophic lateral sclerosis and spinal muscular  
 CC atrophy. IFN nucleic acids and polypeptides are also useful for  
 CC identifying interferon-like proteins and interferon agonists, for  
 CC screening drugs and compounds which inhibit or enhance IFN activity or  
 CC function and as targets for the identification of small molecules that  
 CC are immunostimulatory, immunosuppressive, or stimulate or suppress normal  
 CC cell or tumour cell growth in mammals, including humans  
 XX SQ Sequence 304 AA;  
 Query Match 20.6%; Score 1093; DB 4; Length 304;  
 Best Local Similarity 100.0%; Pred. No. 6e-76;  
 Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 41 EVKEQIORTVATSGNMQAVELLISTLEKGWHLGWTREFVEALRRTGSPLAARYMNPELT 100  
 Db 87 EVKEQIORTVATSGNMQAVELLISTLEKGWHLGWTREFVEALRRTGSPLAARYMNPELT 146  
 QY 101 DLPSPSFENAHDEYLQLLNLQOPTLVDKCMEEELLTTEDRNRRIAAENNGN 160  
 Db 147 DLPSPSFENAHDEYLQLLNLQOPTLVDKCMEEELLTTEDRNRRIAAENNGN 206  
 QY 161 ESGVRELLKRIYQKENWFSAFNLVLRQTNNELVQELTGSDCSENAEENLSQLVQDGQV 220  
 Db 207 ESGVRELLKRIYQKENWFSAFNLVLRQTNNELVQELTGSDCSENAEENLSQLVQDGQV 266  
 QY 221 EEQLLSTTVQPNLEKEVWGMENNSSESSFADSSVVS 256  
 Db 267 EEQLLSTTVQPNLEKEVWGMENNSSESSFADSSVVS 302  
 SQ Sequence 308 AA;

RESULT 14  
 ID AAU00298 standard; protein; 308 AA.  
 XX AAU00298;  
 XX AC;  
 XX DT 12-SEP-2001 (First entry)  
 XX DE Interferon induced polypeptide, IFN6.  
 XX KW Interferon induced nucleic acid; autoimmune disease; lupus erythematosus;  
 KW immunodeficiency; stroke; acquired immunodeficiency syndrome; AIDS;  
 KW graft rejection; viral infection; hepatitis; aplastic anaemia; cancer;  
 KW human immunodeficiency virus; HIV; immune-mediated glomerulonephritis;  
 KW haematologic disease; chronic neutropenia; myocardial infarction;  
 KW neurological disease; Alzheimer's disease; Parkinson's disease; tumour;  
 KW amyotrophic lateral sclerosis; spinal muscular atrophy; human; IFN6.  
 XX OS Homo sapiens.  
 XX PN WO200118208-A2.  
 XX PR 08-SEP-1999; 99US-0152921P.  
 XX PR 20-OCT-1999; 99US-0160575P.  
 XX PR 20-JAN-2000; 2000US-0177104P.  
 XX PR 07-SEP-2000; 2000US-00656633.  
 XX PR 08-SEP-2000; 2000WO-US024704.  
 XX PR 08-SEP-1999; 99US-0152921P.  
 XX PR 20-OCT-1999; 99US-0160575P.  
 XX PR 20-JAN-2000; 2000US-0177104P.  
 XX PR 07-SEP-2000; 2000US-00656633.  
 XX PA (CURA-) CURAGEN CORP.  
 XX PA (BIOJ ) BIOGEN INC.  
 XX PI Peyman JA, Da Silva A, Hochman P, Hsu A;  
 XX WPI; 2001-235201/24.  
 DR N-PSDB; AAS01151.  
 XX PS Claim 1; Page 40-42; 134pp; English.  
 XX PR 08-SEP-1999; 99US-0152921P.  
 XX PR 20-OCT-1999; 99US-0160575P.  
 XX PR 20-JAN-2000; 2000US-0177104P.  
 XX PR 07-SEP-2000; 2000US-00656633.  
 XX PA (CURA-) CURAGEN CORP.  
 XX PA (BIOJ ) BIOGEN INC.  
 XX PI Peyman JA, Da Silva A, Hochman P, Hsu A;  
 XX WPI; 2001-235201/24.  
 DR N-PSDB; AAS01151.  
 XX PS Claim 1; Page 40-42; 134pp; English.  
 XX The sequence represents interferon induced polypeptide, IFN6. IFN nucleic  
 CC acids and polypeptides are useful for treating or preventing a pathology  
 CC associated with IFN polypeptide in a human. They are useful for  
 CC determining the presence of or predisposition to a disease associated  
 CC with altered levels of IFN polypeptide or polynucleotide. IFN nucleic  
 CC acids, polypeptides and antibodies are useful for diagnosis, prevention  
 CC or treatment of variety of immunological and cell proliferative  
 CC disorders, such as autoimmune diseases such as acquired immunodeficiency syndrome  
 CC (AIDS), graft rejection, viral infections including hepatitis and human  
 CC immunodeficiency virus (HIV), immune-mediated glomerulonephritis,  
 CC haematologic diseases such as aplastic anaemia and chronic neutropenia  
 CC and cancer. In addition they are also useful for treating or diagnosing  
 CC various disorders associated with cell death, including myocardial  
 CC infarction, stroke, neurological diseases including Alzheimer's and  
 CC Parkinson's diseases, amyotrophic lateral sclerosis and spinal muscular  
 CC atrophy. IFN nucleic acids and polypeptides are also useful for  
 CC identifying interferon-like proteins and interferon agonists, for  
 CC screening drugs and compounds which inhibit or enhance IFN activity or  
 CC function and as targets for the identification of small molecules that  
 CC are immunostimulatory, immunosuppressive, or stimulate or suppress normal  
 CC cell or tumour cell growth in mammals, including humans  
 XX SQ Sequence 304 AA;  
 Query Match 20.6%; Score 1093; DB 4; Length 304;  
 Best Local Similarity 100.0%; Pred. No. 6e-76;  
 Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 41 EVKEQIORTVATSGNMQAVELLISTLEKGWHLGWTREFVEALRRTGSPLAARYMNPELT 100  
 Db 87 EVKEQIORTVATSGNMQAVELLISTLEKGWHLGWTREFVEALRRTGSPLAARYMNPELT 146  
 QY 101 DLPSPSFENAHDEYLQLLNLQOPTLVDKCMEEELLTTEDRNRRIAAENNGN 160  
 Db 147 DLPSPSFENAHDEYLQLLNLQOPTLVDKCMEEELLTTEDRNRRIAAENNGN 206  
 QY 161 ESGVRELLKRIYQKENWFSAFNLVLRQTNNELVQELTGSDCSENAEENLSQLVQDGQV 220  
 Db 207 ESGVRELLKRIYQKENWFSAFNLVLRQTNNELVQELTGSDCSENAEENLSQLVQDGQV 266  
 QY 221 EEQLLSTTVQPNLEKEVWGMENNSSESSFADSSVVS 256  
 Db 267 EEQLLSTTVQPNLEKEVWGMENNSSESSFADSSVVS 302  
 SQ Sequence 308 AA;

Query Match 20.6%; Score 1093; DB 4; Length 308;  
 Best Local Similarity 100.0%; Pred. No. 6.1e-76;  
 Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 EVKEQIORTVATSGNMQAVELLISTLEKGWHLGWTREFEALRRRTGSPLAARYMNPETL 100  
 91 EVKEQIORTVATSGNMQAVELLISTLEKGWHLGWTREFEALRRRTGSPLAARYMNPETL 150

Db 101 DLPSPSFENAHDEYLQQLNNLQOPTLYDKLVRDVLDKCMEEELLTTIEDRNRRIAAAENNN 160  
 151 DLPSPSFENAHDEYLQQLNNLQOPTLVDKLVRDVLDKCMEEELLTTIEDRNRRIAAAENNN 210

QY 161 ESGVRELLKRIVQKENWFAFLNVLRQTNENLQVQELTGSDCSESNAAEENLSQVQGPQV 220  
 Db 211 ESGVRELLKRIVQKENWFAFLNVLRQTNENLQVQELTGSDCSESNAAEENLSQVQGPQV 270

QY 221 EEQLLSTTVQPNLKEVWGMENNSSESSFADSSVVS 256  
 Db 271 EEQLLSTTVQPNLKEVWGMENNSSESSFADSSVVS 306

Query Match 16.2%; Score 861; DB 4; Length 267;  
 Best Local Similarity 96.6%; Pred. No. 4.7e-58;  
 Matches 170; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 41 EVKEQIORTVATSGNMQAVELLSTLEKGWHLGWTREFEALRRRTGSPLAARYMNPETL 100  
 Db 87 EVKEQIORTVATSGNMQAVELLSTLEKGWHLGWTREFEALRRRTGSPLAARYMNPETL 146

QY 101 DLPSPSFENAHDEYLQQLNNLQOPTLVDKLVRDVLDKCMEEELLTTIEDRNRRIAAAENNN 160  
 Db 147 DLPSPSFENAHDEYLQQLNNLQOPTLVDKLVRDVLDKCMEEELLTTIEDRNRRIAAAENNN 206

QY 161 ESGVRELLKRIVQKENWFAFLNVLRQTNENLQVQELTGSDCSESNAAEENLSQVQGPQV 216  
 Db 207 ESGVRELLKRIVQKENWFAFLNVLRQTNENLQVQELTGSDCSESNAGICNTEED 262

Search completed: June 2, 2004, 19:10:14  
 Job time : 67 secs

RESULT 15  
 AAU00297 standard; protein; 267 AA.

ID AAU00297  
 XX  
 AC AAU00297;  
 XX DT 12-SEP-2001 (first entry)  
 XX DE Interferon induced polypeptide, IFN5.  
 XX Interferon induced nucleic acid; autoimmune disease; lupus erythematosus;  
 XX immunodeficiency; stroke; acquired immunodeficiency syndrome; AIDS;  
 XX graft rejection; viral infection; hepatitis; aplastic anaemia; cancer;  
 XX human immunodeficiency virus; HRV; immune-mediated glomerulonephritis;  
 XX haematologic disease; chronic neutropenia; myocardial infarction;  
 XX neurological disease; Alzheimer's disease; Parkinson's disease; tumour;  
 XX amytrophic lateral sclerosis; spinal muscular atrophy; human; IFN5.  
 XX Homo sapiens.  
 XX PN WO200118208-A2.  
 XX PD 15-MAR-2001.  
 XX PF 08-SEP-2000; 2000WO-US024704.  
 XX PR 08-SEP-1999; 99US-0152921P.  
 XX PR 20-OCT-1999; 99US-0160575P.  
 XX PR 20-JAN-2000; 2000US-0177104P.  
 XX PR 07-SEP-2000; 2000US-00656633.  
 XX PA (CURA-) CURAGEN CORP.  
 PA (BIOJ ) BIOGEN INC.  
 XX PI Peyman JA, Da Silva A, Hochman P, Hsu A;  
 XX DR WPI; 2001-235201/24.  
 XX DR N-PSDB; AAS01150.  
 XX New interferon induced polypeptides and polynucleotides, useful for the  
 PT diagnosis, prevention and treatment of immunological, cell proliferative  
 PT disorders, such as lupus erythematosus, cancer, stroke and Alzheimer's  
 PT disease.  
 XX Claim 1; Page 33-35; 134pp; English.  
 XX The sequence represents interferon induced polypeptide, IFN5. IFN nucleic  
 CC acids and polypeptides are useful for treating or preventing a pathology  
 CC associated with IFN polypeptide in a human. They are useful for  
 CC determining the presence of or predisposition to a disease associated  
 CC with altered levels of IFN polypeptide or polynucleotide. IFN nucleic  
 CC acids, polypeptides and antibodies are useful for diagnosis, prevention

GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2004, 19:08:58 ; Search time 23 Seconds  
 (without alignments)  
 2300.723 Million cell updates/sec

Title: US-09-515-363C-2  
 Perfect score: 5311  
 Sequence: 1 MSNGYSTDENFRYLISCPRA.....LPITFPNLDYSECCLESDED 1025

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database :

Issued Patents AA: \*  
 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep: \*  
 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep: \*  
 3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep: \*  
 4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep: \*  
 5: /cgn2\_6/ptodata/2/iaa/PCITS\_COMB.pep: \*  
 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	286	5.4	503	4	US-09-408-020-66	Sequence 66, Appl
2	284	5.3	502	4	US-09-408-020-34	Sequence 34, Appl
3	217.5	4.1	2662	4	US-09-595-684B-31	Sequence 31, Appl
4	202	3.8	3878	4	US-09-914-259-11	Sequence 11, Appl
5	194.5	3.7	2482	1	US-08-328-254-6	Sequence 6, Appl
6	193.5	3.6	661	4	US-09-107-532A-3677	Sequence 3677, Appl
7	191.5	3.6	3248	1	US-08-353-700-1	Sequence 1, Appl
8	191.5	3.6	3248	5	PCT-US95-16216-1	Sequence 1, Appl
9	188.5	3.5	1388	2	US-08-685-576-4	Sequence 4, Appl
10	188.5	3.5	1388	4	US-09-976-594-296	Sequence 296, Appl
11	187.5	3.5	2125	4	US-09-919-172-29	Sequence 29, Appl
12	186.5	3.5	1786	3	US-08-973-462-8	Sequence 8, Appl
13	182.5	3.4	2954	4	US-09-150-867-1	Sequence 1, Appl
14	179	3.4	666	4	US-09-134-001C-5465	Sequence 5465, Appl
15	174	3.3	1388	2	US-08-685-576-1	Sequence 1, Appl
16	173	3.3	1211	4	US-09-134-001C-4820	Sequence 4820, Appl
17	171.5	3.2	976	3	US-09-104-324B-4	Sequence 4, Appl
18	171.5	3.2	2285	4	US-09-308-375-2	Sequence 2, Appl
19	167.5	3.2	1219	4	US-09-107-532A-6020	Sequence 6020, Appl
20	167	3.1	677	4	US-09-328-352-4365	Sequence 4365, Appl
21	167	3.1	3696	4	US-09-134-001C-5080	Sequence 5080, Appl
22	166.5	3.1	1886	4	US-08-938-105-3	Sequence 3, Appl
23	164	3.1	1354	3	US-08-685-871-2	Sequence 2, Appl
24	163	3.1	1939	4	US-09-310-187A-1	Sequence 1, Appl
25	159.5	3.0	956	4	US-09-134-001C-4452	Sequence 4452, Appl
26	159	3.0	442	4	US-09-489-039A-1177	Sequence 11770, Appl
27	158	3.0	781	2	US-08-675-631-3	Sequence 3, Appl

#### ALIGNMENTS

RESULT 1	US-09-408-020-66	;	Sequence 66, Application US/09408020
	;	Patent No. 6632937	
	;	GENERAL INFORMATION:	
	;	APPLICANT: Swanson, Ronald V.	
	;	;	TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAUM SYMBIOSUM
	;	;	FILE REFERENCE: DCORP .002A
	;	;	CURRENT APPLICATION NUMBER: US/09/408,020
	;	;	CURRENT FILING DATE: 1999-09-29
	;	;	PRIOR APPLICATION NUMBER: 60/102,294
	;	;	PRIOR FILING DATE: 1998-09-29
	;	;	NUMBER OF SEQ ID NOS: 123
	;	;	SOFTWARE: FastSEQ For Windows Version 3.0
	;	;	SEQ ID NO 66
	;	;	LENGTH: 503
	;	;	TYPE: PRT
	;	;	ORGANISM: Cenarchaeum symbiosum
	;	;	US-09-408-020-66

Query 655 DGDEDDEDDLKKPKLKDDETDRFLMNLFFENNKKMLKRLAENPEYENELTKLRLNRTIMEQYTR 714 Db 342 -RGARGKALVFTSYRDSVDLI-----HSRLKAAGINSGLLIGKAGE---KGLKORKQ 389  
 Db 307 -----LFEEDR----NFTGAMARAKAAQAGMEEHPKPKL-----337 Qy 772 KEVISKFRTGKINLLIATTAEGLDIKECNIVIRYGLVTNEIAMVOAGRGR-ARADESTY 830  
 Query 715 TEESARG----IIFTKTRQSAVALSQWITENEKFAEVGVKAHHLIG-AGHSSEFKPMTO 768 Db 390 VETVAKFRDGGYDVLVSTRVGEEGLDISEVNLVIFYDNPSSIRYVQRRGRTGRKDAGRL 449  
 Db 338 -EEAVRGAKGKALVFTSYRDSVDL-----HSKLQAAGINSGLIGKAGE---KGLKQ 386 Qy 831 VLVAHSGS 838  
 Db 769 NEOKEVISKEPRTGKINLLIATTAEGLDIKECNIVIRYGLVTNEIAMVOAGRGR-ARADE 827 Db 450 IVLMAGKT 457  
 Db 387 KKOVETVAKFRDGGYDVLVSTRVGEEGLDISEVNLVVFYDNPSSIRYVQRRGRTGRKDA 446  
 Query 828 STYVLVAHSGS 838 Db 447 GKLVVLMAKGT 457  
 Db 447 GKLVVLMAKGT 457  
 RESULT 3  
 US-09-595-684B-31  
 Sequence 31, Application US/09595684B  
 Patent No. 6544766  
 GENERAL INFORMATION:  
 / APPLICANT: Beraud, Christophe  
 / APPLICANT: Ohashi, Cara  
 / APPLICANT: Sakowicz, Roman  
 / APPLICANT: Vaisberg, Eugeni  
 / APPLICANT: Wood, Kenneth  
 / APPLICANT: Yu, Ming  
 / TITLE OF INVENTION: Human kinesins and methods of producing  
 / TITLE OF INVENTION: and purifying human kinesins  
 / FILE REFERENCE: cytop036  
 / CURRENT APPLICATION NUMBER: US/09/595,684B  
 / CURRENT FILING DATE: 2002-06-24  
 / PRIOR APPLICATION NUMBER: 09/295,612  
 / PRIOR FILING DATE: 2000-04-20  
 / NUMBER OF SEQ ID NOS: 105  
 / SOFTWARE: FastSEQ for Windows Version 4.0  
 / SEQ ID NO 31  
 / LENGTH: 2662  
 / TYPE: PRT  
 / ORGANISM: Human  
 US-09-595-684B-31  
 Query Match 4.1%; Score 217.5; DB 4; Length 2662;  
 Best Local Similarity 18.6%; Pred. No. 7.6e-09;  
 Matches 209; Conservative 185; Mismatches 421; Indels 307; Gaps 47;  
 Qy 9 ENFRYLISCFRARVVKMYIQVEPVLDYLTLFLPAEVKEQIORTVATSGNMQAVELLSTLEK 68  
 Db 1039 EQQRKIFSLIQEKNLQQMLVESVIAKEQLKTDKENIEMTIE---NQEERLRLGDELKK 1095  
 Qy 69 GWNHLGWTREFFV----EALRRTGSELAAARYMNPETLDLPSPSFENAHEDEYLQOLLNLQP 123  
 Db 1096 -----QQEIVAQXNHAIKEGELSRCDLAEVEEKLKEKSQQLQEQQQQLNQOEE 1148  
 Qy 124 TLVDKLLVRYDV--LDKCMEEELLIE--DRNRLAAAEN-NGNESGVRELLK-RIVQKE-- 175  
 Db 1149 MSEMQKKINEENLKNKELTLEHMETERLEAQKLNEYEEVKSITKVKVILKELQ 1208  
 Qy 176 -----NWESAFNLVLRQTG-----NNELVQELTGSDCSESNAEIELN 212  
 Db 1209 KSFETERDHLRGYIRETEATGLQTKELKIAHHLKEHQETIDELRRS-VSEKTAQINT 1267  
 Qy 213 SQVDPQV----EEQLLSTTVQPNLEKEVWGMENNSSESSEADSSVVSSESDTS 261  
 Db 1268 QDLIEKSHTKLQEEITPVLHEEQELL----PNVKKVSETQETMNEELLTEQSTTKDS-TT 1321  
 Qy 262 LAEGSVSCLDESLGHNSNMGSDEENVAARASPEPELQLRPYQMEVAQPALE 321  
 Db 1322 LARIEM----ERLRLNEKF-----QESQEEIKSLTKERDNLXKTIKEALEVKHDQL- 1367  
 Qy 322 GRNIIICLPTGSGKTRVAVYIAK-----DHLDDKKKKASEPGKVVVLVNLVLEQL 372  
 Db 1368 -----KEHIRETLAKIQEKSQKQESLNMKEKDNNETTKI-----VSEMEQ- 1407  
 Qy 373 FRKEFQPFLRKWYRVTGIGSGDTQKLISFPE---VVVKSCDIIISTAQILNLSLNLENG-- 427  
 Db 1408 FKPDSALRIELEMGLIS--KRLQESHDMEKSVAKERDDIQLQEVIQSESQDQIKENIK 1465  
 Query 715 TEESARG----IIFTKTRQSAVALSQWITENEKFAEVGVKAHHLIG-AGHSSEFKPMTO 771

428	EDAGVQLSDFLSLIITIDECHHTNKEAVYNNIMRHYLMQKLKNNRLLKENKPNVPLPQILGL	4.87	QY	57	QAVELLJLSTLEKGWHLGWTREFEALRRGSPSPLAARYMNPEL-TDLPLSPSPFENAHDDEYL	115
1466	EIAVAKHLETEEELKVAHCCLKEQETIN-----ELRVNLSEKETE-----	15.05	Db	1194	KPLHLLIGKLQAV-----SEECSYFLOTLCSVLGEYYPALKCEVNAEDKENSGD--	
488	TASPGVGGATKOAKA-----EEHILKLCANLDAFTIKTYKENLDOLKNQIQEPCKKF	53.9	QY	116	QLNNLQOPTLVD-KLLYRD---VLDKCMBE--ELLTIEDR-NRIAAEENNNG--ES 162	
1506	-----ISTIQKOLEAINDKLQNQIQEYKEEQLNIKQISEVQENVNELK-QFKEHRK--	15.57	Db	1245	YISENEDPELQDYEVQDFQENMHTLLNKVTEYNNKLVLQTRLSKWIWGOQTDMKLF	1304
540	AIAADATREDPFKEKLLIMTRIQTYCQ--MSPMSDFTGTOPYEWAQMEKKAAKGNR--	5.95	QY	163	GYRELLJRKIVOKENWESAFNLNVJLQRTGNNE-----LYQELTGSDCSESNAEIEN	211
1558	-----AKDSALQSIESKMLLETNRLQESQEEQIOMIKEKEEMRVRQEAQJIERDQLKENTKEI	16.15	Db	1305	GREENLPKEETE-----PLSIHSQMTNLIEDIDVNHKSKLSSLQDLEKTKLKEEQVQELES	1357
596	-----KERVVC-AEHLRKYNELLQIN-----DTIIRMID-AYT	6.24	QY	212	LSQVDGPQVEEQLLSTVOPNLKEVWGMENN--SSESSF----ADSSVVSSESDTSIA	263
1616	VAKMKESQEKEYQFLKMTAVNETQERMCIEIEHLIKEQFETQKLNLLENTEENIRLTLQILHE	16.75	Db	1358	LIS---SLQQQLKET--EQNYPEAEIHCLQKRLQAVSESTVPPSLPVDSVVITESDAQRT	1411
625	HLETFTYNEEKDKKFAVIEDDSDEGGDEYCDGDEDDEDDLKPKLKD-----	6.70	QY	264	EGSVSCULDESGLHNNSNMCGSDGTMGSDDEENVAARASPEPELQLRPYQMEVAQPALEGK	323
1676	NLEEMRSVTKE-----DDLRSVEETLKVERDQLKENLRETITRDLEXQEEULKIVHMHL	17.29	Db	1412	MYPGSCVKKNTIDGTIEFSGEFFGV---KEETNIV-----KLEKQYQEQQLEE---EVA	1457
671	-----ETDRFLMTLFFENNKMVLKRLAENPEYENKL-----TKLRNRTIMEQYTRTE	7.16	QY	324	NIITCLP-----TGSGKTRVAVYIAKDHDLKDKKKASEPGKIVLVMK	365
1730	KEHQETIDKLKGIVSEKTNBEISNMQKDLHESNDALKADQDLKIQEEELRIAHMHLKEQQETI	17.89	Db	1458	KVIVSMSIAFAAQOTEELSRSRISGGKENTASSKQAHAVCQQEQHYFNEMKLSQQ---IGFQT	1514
717	ESARGIIIFTKTRQAYALSQWITEN---EKFPAEVGVKAHHLIGAGHS-----SEFK	7.64	QY	366	VLLVEQLFRKEFQFPLKKWYRVLGSG-DTQLKTKISFPEVV--KSC-----407	
1790	DKLRGIVSEKTDKLSNMQKDLENSNAKLOQEKIQLKANEHQLITLKKDVNETQKVSOME	18.49	Db	1515	FETVDPVKEEKFPLSKKE-----LGEHGKEILLISNSDHPIDESKDCVLTISEEMFSKDKT	1570
765	PMTOQ--NEQKEVISKFRIGKINLLIATTVAEGLDLIKECNIVIRGLVLTNEIAMQARGR	8.22	QY	408	-----DIIISTAQILENSLLNLENEDAGVQL-----SD 436	
1850	QLRKQIKDQSLTTSKLEIENLNL--AQLHENLIEEMK-----SYMKERDN	18.92	Db	1571	FIVROQSHDEISVSMASRQLMNEEQLEDMRMQLNEEQLDMRQELVSE	1630
823	ARADESTYVLAHSGSGV-----EEH-ETVNDFREKOMYKAH 85.9		QY	437	FSLTIDE-----CHHTNKEAVYNN-----456	
1893	LRRVEETLKLERDQLKESLQETKARDLEIQQEQLKTARMLSKEHKEVTDKLREKISEKTIQ	19.52	Db	1631	RERVILLEEALKQLSLAGREKLUCCELRNSSTOTONGNENQGEVEEQTFKEKEKELDRKPED	1690
860	CVQNMK-----PEEYAHKKILELOMOSIM-----ERRMKTKRNIAKHYKNPNSLI	9.03	QY	457	-----IMRHYLMQKLKNQNLKRNQKPKVTPLPQIQLGLTASPVGFG-----495	
1953	ISDIQKDLDSKSKDELQKKIQLQELQKKEVTDKLREKISEKTIQ	19.52	Db	1691	VPPEILSNERYALQK-ANNRLLKILLEVVKTTAAVEETIGRHLVGLILDRSSKSQSSASLI	1749
904	TEFLCKNCVLAACSGEDIHIVTEKMHVNMTPEFKELYIREN 94.5		QY	496	-----ATKQAKAEEHILKLCANLDRAFTIKTVKENLDOLKNQIQLEPCKK-----F 539	
2009	-YLCK-----CEMDNFQLTKLHE-----SLEEIRIVAKER	20.08	Db	1750	WRSEAAEVVKSCVHEHTRTIDESIPSYGSGDMPRNDINMWSKVTEEGTTELSQLVRSGF	1809
4	ISU		QY	540	AIADATREDPEKEKL-LEIMTRIQTYCQ--MSPMSDFGTQ---PYBQWAJQMEKKAQK	592
6-09-914-259-11	Sequence 11, Application US/09914259		Db	1810	A--GTEIDPENEELMLNINSSRLQAAVEKLLAISETSSQLEHAKVTQTELRESFRQKQ	1866
	Patent No. 6495336		QY	593	GNRKERVCAEHLRK--YNE-----ALQINDTIRMDAYTHLETFYNEEKDKKFAVIT--	641
	APPLICANT: Makowski, Lee		Db	1867	EATESLKCQEEELRERLHEESRAREQLAVELSKAEGVTDGYADEKTLFERQIEKTDIIDR	1926
	APPLICANT: Hyman, Paul		QY	642	-----EDDSDEGGDE-----652	
	APPLICANT: Williams, Mark		Db	1927	LEQELLCASNRLQEELEAQQQIQEEERLLSRQKEAMKAEGPVEQOLLQETEKLKKEKLE	1986
	TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES		QY	653	-YCDGDEDDEDDLKKPLKLDDET-----RFLMLTLEFFENKMLKRLAENPEYENELTKLRL	705
	FILE REFERENCE: 8471-010-999		Db	1987	706 VQCQAEKVDDLQKVKALEIIVDEEQVSRFI-ELEQEKTTELMDLROQONQALEKQLERMR	2045
	CURRENT APPLICATION NUMBER: US/09/914, 259		QY	706	NTIMEQYTRTEESARGLIFTKTROSAYALSQWITENEKEFAEVGVKAHLLIGAGHSSEFPK	765
	CURRENT FILING DATE: 2000-11-21		Db	2046	KFLDEQALDREHERD-----VFOQEIQKLEQOLKV-----VPRFQFOP	2081
	NUMBER OF SEQ ID NOS: 180		QY	766	MTQNEQKEVIS-----KFRTGKINLLIATVAAEGPVEGL--DIKECNIVIRGLVLTNEIAMVQ	818
	SOFTWARE: FastSEQ for Windows Version 4.0		Db	2082	819 ARGRARADESTYVLAHSGS-----GVIE-----	842
	SEQ ID NO: 11		QY	2130	2130 FRVR----ELEQALVSADETFQKVEDRKHGFAVEAKPELSLEVQQAERDAIDRKEKEIT	2185
	LENGTH: 3878		QY		-----HEVNDREKMMYKAIHCVCNMKPEEYAHKILELOMOSIMEKMTKTR-----NIAKH	893
	TYPE: PRT		QY		-----KET	
	ORGANISM: Homo sapiens		QY			
3-09-914-259-11	Query Match	3.8% ; Score 202 ; DB 4 ; Length 3878 ;	Db			
	Best Local Similarity	18.8% ; Pred. No. 3e-07 ; Mismatches 424 ; Indels 428 ; Gaps 55 ;	Db			
	Matches 238 ; Conservative		Db			
	5 YSTD-----ENFYLISCFRARVKMYIQVEPVLDYLTLPAEVKEQIORTVATSGNM	56	QY			
	11136 YSTHVQDQREYMEENEKDKAICSLKEELIFIAQEBKIKELQKIHQEL--QTMKTOETGDEG	11.93	Db			

RESULT 5

US-08-328-254-6

Sequence 6, Application US/08328254

Patent No. 5710022

GENERAL INFORMATION:

APPLICANT: Zhu, Xueliang

TITLE OF INVENTION: A No. 5710022el Nuclear Mitotic Phosphoprotein

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Campbell and Flores

STREET: 4370 La Jolla Village Drive, Suite 700

CITY: San Diego

STATE: California

COUNTRY: USA

ZIP: 92122

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US 08/328,254

FILING DATE: 24-OCT-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/141,239

FILING DATE: 22-OCT-1993

ATTORNEY/AGENT INFORMATION:

NAME: Campbell, Cathryn A.

REGISTRATION NUMBER: 31,815

REFERENCE/DOCKET NUMBER: P-CJ 1191

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 535-9001

TELEFAX: (619) 535-8949

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 2482 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-328-254-6

Query Match 3.7%; Score 194.5; DB 1; Length 2482;

Best Local Similarity 19.2%; Pred. No. 6.5e-07;

Matches 219; Conservative 186; Mismatches 473; Indels 265; Gaps 46;

QY 21 RVKMYIQVEPVL--DYLTFLPAEVKEQIQTATSG----NMQAVELLIS 64

Db 1189 KIEACIELEKTVGELKKENSDLSKLEYFSCDHQELLQRVETSEGNSDLEMADKSSRE 1248

QY 65 TLEKGWHL--GWTRFVEALRRRTGSPLAARYMNPELTDLPSPSFENAHDDEYLQILLNLLQ 122

Db 1249 DIGDNVAKVNDSWKERFLD-----VENELSRIRSEKASIEH-EAILYLEADLE 1294

QY 123 PTLVDKL-LYRVDVLDK----CMEFEELTTI-EDRNRIAAENNGNE--SGVRELLKRIQ 173

Db 1295 VVQTEKLCLEKDNEKQKVIVCLEEELSVVTSERNOLRGELDTMSKTTALDQLSKEMKE 1354

QY 174 K---ENWESAFLNVLROT----GNNELVQELTGSDCSESNAELENLSQ-VDGPQVEEQ 223

Db 1355 KTQELESHOSECLHCIQVAEAEVKEKTELLQTLIS-SDVSELLIKDKTHQEKLSLEKDSQ 1413

QY 224 LLSST-----VQPNLEKEVWGMENN-----SESSFADESSVVSSESDTSLAEGS---- 266

Db 1414 ALSLTCKCELENQIAQLNKEKEKELLVKESESLOQARLSESDYEKLNVSKALEAALVEKGEFAL 1473

QY 267 -VSCLDESLGHNSNMGSDSGTMGSDE-----ENVAARASPEPELQLRPYQMEVVAQP 318

Db 1474 RLSSTQEEV-HQLRRGIEKLRVRIEADERKQHIAEKLRERENDSLKDVKENLRELQ 1532

QY 319 ALEGKNNIICLPTGSGKTRAVYIAK-DHLDKKKKASEPCKVIVLNKVLLVEQLERKEF 377

Db 1533 MSEBNQEVILDAENSKAIEVETLKTQIEEMARSILKVFELDLVTLRSEKENLTKQIQEKQG 1592

QY 378 QPFLKWKWYRVIQ-----LSGDTQLKISFPEVVKSCDIIISTAQILENSILNLEN---- 426

Db 1593 Q--LSELDKLSSFKSLLKEKEQAEIQIKEESK-----TAVENLQNQNLKELNEAVAAALC 1644

QY 427 GEDAGVQLSDFLIIIDECHTNKEAVYNNIMRHYLMQKLKNNRKEENKPVIPLEQI-- 484

Db 1645 GDQELEMKATEQSLDPPPIEEHQLRNSI-----EKLRARLEADEKKQLCQLVQLIKE 1694

QY 485 -----LGLTASPGVGGATKQAKAEEHILKL----CANLD 514

Db 1695 SEHHADLLKGRENLERELLEARTNQEALEAENSKGEGEVELKAKIEGMTQSLRGLELD 1754

QY 515 AFTIKTKVENLDQKNOIQEPCKKFAIADATREDPFKEKLIEMTRIOTYQMSPMDSDFG 574

Db 1755 VVTJIRSEKENLTLQKBERISELELISSFENTILOEK-----EQEKVQMKERKSSTA 1807

QY 575 TQPYEQWAIQMEKKAAKKGKGNRKERVCA--EHLRKYNEALQIN--DTIRIMIDA---YTHL 626

Db 1808 MEMLQTLQKLKEENRVAALHNDQEACKAKEQNLSSQVCECLEKAQOLLQGLDEAKNNYIVL 1867

QY 627 ETFVN-----EEKDKKFAVIEDDSDEGGDEYCDGDEDDEDLKKPLKLDDETD 673

Db 1868 QSSVNGLIOQEVEDGKQKLEKKODEEISRLKQNI-----QDQEQLVSKLSQVEGE 1915

QY 674 RFLMTLFFENNKKMLKRLAENPEY-----ENEKLTKLRNNTIMEQYTRTEESARGLIIFTK 726

Db 1916 H--QLWKEQNLLELRNLTVLEOKIQVLOQSKNASLQDTLEVQSSSYRNLENE--LELTK 1969

QY 727 TRQSAYALSQWITENEKPAEVGVKAHHLIGACHSSEFKPMTQNEQKEVISKFTGKINLL 786

Db 1970 MDKMSFV-----EKVNQMTAKETELQREMHEMAQKTAELQEEFSGEKNRLAGELQJL 2021

QY 787 -----IATTVAEEGLDIKECNCNTIVTRYGLVTNEIAMYQARGRARADESTYV 831

Db 2022 LEEIKKSSKDQKLKELTLENSELKKSSLDCMHKDQVEKEGKVREIAEYQLRHEAKKHOAL 2081

QY 832 LVAHSGSGVIEHETVNDFREKMYKAIIHCQVNMMPE-----EYAHKILELQMQSIME 883

Db 2082 LLDTPNKQYEVETQ---YREKLTSKKE-LCLSSQKLEIDLLKSSKEELNNSLKATTQ-ILE 2136

QY 884 KKMFTTK-----RNTAKHYKNNPNSLITFLCKNCVLAQSGEDIHYTEK----- 925

Db 2137 ELKKTKMDNLYKVNQNLKKEKNERAOGKMKLLIKSKQLE---EEKEILQKELSQLQAAQEK 2193

QY 926 -----MHHVNMTPEFKELYIRENKA-----LQKKCADYQINGELICKCGQAWGTM 971

Db 2194 QKTRGTVMDTKVDELTTEIKELKETLEKTEADEYLDKYCS-LLISHEKLEKAKMLETQ 2252

QY 972 MVH 974

Db 2253 VAH 2255

RESULT 6

US-09-107-532A-3677

Sequence 3677, Application US/09107532A

Patent No. 6583275

GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

NUMBER OF SEQUENCES: 7310

CORRESPONDENCE ADDRESS:

ADDRESSEE: GENOME THERAPEUTICS CORPORATION

STREET: 100 Beaver Street

CITY: Waltham

STATE: Massachusetts

COUNTRY: USA

ZIP: 02454

COMPUTER READABLE FORM:

MEDIUM TYPE: CD/ROM ISO9660

COMPUTER: PC

OPERATING SYSTEM: <Unknown>

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/107,532A

FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/085,598

FILING DATE: 14 May 1998

APPLICATION NUMBER: 60/051571

FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Ariniello, Pamela Deneke

HYPOTHETICAL: YES

REGISTRATION NUMBER: 40,489

REFERENCE/DOCKET NUMBER: GTC-012

TELECOMMUNICATION INFORMATION:

TELEPHONE: (781) 893-5007

TELEFAX: (781) 893-8277

INFORMATION FOR SEQ ID NO: 3677:

SEQUENCE CHARACTERISTICS:

LENGTH: 661 amino acids

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: YES

ORIGINAL SOURCE:

ORGANISM: Enterococcus faecium

FEATURE:

NAME/KEY: misc feature

LOCATION: (B) LOCATION 1..661

SEQUENCE DESCRIPTION: SEQ ID NO: 3677:

US-09-107-532A-3677

Query Match 3.6%; Score 193.5; DB 4; Length 661;  
Best Local Similarity 18.8%; Pred. No. 9.4e-08;  
Matches 147; Conservative 130; Mismatches 240; Indels 267; Gaps 36;

Query 299 RASPEPELQLRPyQMEVAQP-----ALEGKNTIICL-PTGSGKTRVAVYIAKDHLD 348

2 RASRHFEL-VSKYQPGDQPEAINQLVGVGGKAQILLGATGTGKT----YTISNLIE 56

Query 349 KKKKASEPGRKVIVLVIKKVLLQFR--KEFQPLFKWVYRIGLSSGDTQLKISF----P 401

57 KVNKPT----LIIAHNKTLAGQLYGEFKEFFP-----NNAVEYFVSYDDYYQP 100

Query 402 EV-VKSCDIIISTAQILLENSLLMLENGEDAG-VQLSDSLITIIDE-----HHTNKEA 452

101 EAYVPPSSDTYIEKDSSVNDIIDLRRHSATSSLERNDVIVTAVSCTIFGLGSPFPEYQKV 160

Query 453 V-----YNNIMMRHMLMQKLXNN-----RLKKENKPVPLP-----482

161 VSIRQGAELDRNQLIRDVLVSIQFERNDIDFQRGRFRVGDVVEIFPASRDERALRVEFFG 220

Query 483 -----QILGLTASPVGGGATKQAKAEEHILKLCANLDAFTIKTKTRENLDQ 528

Db 221 DEVERIREVNALTGEVLGETEHVIAIFPATHFVTNDEHMEHAVANIKA---ELEQRRLTVL 276

Query 529 KNQIQEPCKKFIAADATREDPFKEKLEIMTRIQ-TCYQMSPMDFG-TQPYEQWAQIOME 586

277 RNE-----NKLLEAQRLEQRTNYDIEMMLEMGTSGIENYSRHMD 316

Query 587 KAAKKGNNRKERVCAEHLRKYNEALQINDTIRNIDAYTHLETYFNEEKDKKFAVIEDDSD 646  
Db 317 GR--KEGE-----PPYTLDD-FFPED---FLIVADESH 343

Query 647 -----KKPLKLDDETDRFLMTLFFENFKML 687  
Db 344 VTMPQIRG---MYNGDRARKQMLVDYGFRPLPSALDNPLRLEEEFKHV----NQII 392

Query 688 KRLAENPEYENK-----LTKLRNNTIMEQYTRTEESARGII 723  
Db 393 YVSATPGPYEHQTDVTIQQQIIRPTGLLDPVIEVRPIMQIDDLVGENERVEKDQRVVF 452

Query 724 FTKTRQSQAYALSQWITENEKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKFRTGKI 783  
Db 453 TTLKKMABEELTDY----FKELGIVKVYL----HSDIKTL---ERTEIIRDRLIGEF 498

Query 784 NLLIATTVAEGLDIKECNIVI----RYGLVLTNEIAMVQARGRARADESTYVLVAHSGS 838  
Db 499 DVLVGINLLREGLDVPEVSLVAILLDAKDEGLPLRSERSLSVQTIGRAARNEEGKVIM---- 553

Query 839 GVIEHETVNDFREKMMYKAIHVQNMKPEEYAHKILLEQMSIMEKKMKTKRNIAKHYKN 898  
Db 554 -----YADKVTD-SMRLAMDETSR-RTTIQKQYNE 581

Query 899 N----PSLITFLCKNCVLAACSGEDIVIEKMHVNMTPEFKELYVRENKALQ--KKC 951  
Db 582 EHGVYPKTLIKEIIRDLSITKESEDDTKEAVQVSYEEMTKEEKDTLIMMCKLEKEMKDAAKA 641

Query 952 ADYQ 955  
Db 642 LDFF 645

RESULT 7  
US-08-353-700-1  
; Sequence 1, Application US/08353700  
; Patent No. 5599919  
; GENERAL INFORMATION:  
; APPLICANT: RATTNER, JEROME B.  
; APPLICANT: YEN, TIMOTHY J.  
; TITLE OF INVENTION: NUCLEIC ACID ENCODING A  
; TITLE OF INVENTION: TRANSIENTLY-EXPRESSED KINETOCHEORE PROTEIN,  
; TITLE OF INVENTION: AND METHODS OF USE  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DANN, DORFMAN, HERRELL AND SKILLMAN  
; STREET: 1601 MARKET STREET, SUITE 720  
; CITY: PHILADELPHIA  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19103-2307  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/353,700  
; FILING DATE: 09-DEC-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: REED, JANET E.  
; REGISTRATION NUMBER: 36,252  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 563-4100  
; TELEFAX: (215) 563-4044  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3248 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein

HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 ORIGINAL SOURCE:  
 ORGANISM: HUMAN  
 US-08-353-700-1

Query Match 3.6%; Score 191.5; DB 1; Length 3248;  
 Best Local Similarity 19.2%; Pred. No. 1.8e-06; Gaps 47;  
 Matches 221; Conservative 184; Mismatches 465; Indels 281; Gaps 47;

QY 21 RVKMYIQVEPVL----DYLTFPLAEVKEQIORTVATSG----NMQAVELLIS 64  
 Db 1917 KIEACIELEKIVGELKKENSDSLSEKLEYFSCDHQELLQRVETSEGLNSDLLEMADKSSRE 1976

QY 65 TLEKGVWHL--GWNTREFVEALRRRTGSPLAARYMNPELTDLPPSPSFENAHEDEYLQLLNLLQ 122  
 Db 1977 DIGDNVAKVNDSWKERFLD----VENELSRIRSEKASIEH-EALYLEADLE 2022

QY 123 PTLVDKLLVRDVLDK----CMEEEELLTI-EDRNRIAAENNGNE--SGVRELLKRIVQ 173  
 Db 2023 VVQTEKLCLEKDNNENKQKVIVCLEEELSUVVTSERNQLERGELDTMSKTTALDQLSEKMK 2082

QY 174 K----ENWFESAFNLNVLRQT----GNNELVQELTGSDCSESNAELENLSQ-VDGPQVEEQ 223  
 Db 2083 KTQELESHQSECLHCIOVAEAEVKEKTELLQTLTS-SDVSELLKDTHQEKLQOSLEKDSQ 2141

QY 224 LLSTT----VQPNLEKEYWGMENNNS----SESSFADESSVYSESDTSLAEGS----266  
 Db 2142 ALSLTKELENNQIAQLNKEELVKESESLSQARLSESVDYEKLNVSKALEAALVKEGFFAL 2201

QY 267 -VSCLDLSELGHNSNMGSGDSGSDSDE----ENVAARASPEPELQLRPYQMEVAQP 318  
 Db 2202 RLSSTQEEV-HQLRRGIEKLRVRIEADEKKQLHIAEKLKERERENDSLSKDKYENLRELQ 2260

QY 319 ALEGKNIILICLPTGSKGKTRVAVYIAK-DHLDKKKKASPEPGKVIVLNKVLLVEQLFRKEF 377  
 Db 2261 MSEENQELVILDAENSKAEEVETLKTQIEEMARSLSK1KFELDLVTLRSEKENLTQ1QERQG 2320

QY 378 QPFKKWYRIG----LGGDTQLKISFPEVVKSCDII1STAQILENSLLNLEN----426  
 Db 2321 Q--LSELDKLSSFKSLLEEKEQAEIQIKEESK----TAVEMLNQNLKELNEAVRALC 2372

QY 427 GEDAGVQLSDFLSLIIDECHHTNKEAVYNNIMRHYLMOQLKNNRLKKENKPV1PLPQI--484  
 Db 2373 GDQEIMKATEQSLDPPIEEHQLRNSI----EKLARLADEKKQOLCVLQQLKE 2422

QY 485 ----LGLTASPGVGATKQAKAEEHILKL----CANLD 514  
 Db 2423 SEHHADLJKGRVENLERELEIARTNOEHALEAENSKGGEVETLKAIEGMTOSLRGLELD 2482

QY 515 AFTIKTVKENLDQLKNOQIOPCKKFAIADATREDPFKEKLLIMTRIQTYCQMSPMDFG 574  
 Db 2483 VVTIRSEKENLTNEQKEQERISELEIINSSFENILOEK----2521

QY 575 TQPYEONAIQMEKKA----KKGNRKERVCAEH----LRKYNEALQIN---614  
 Db 2522 ---EQEKVQMKEKSSTAMEMLQTQKELN--ERVAALHNDQEACKAKEQNLSSQVECLE 2575

QY 615 -DTIIRMIDAYTHLETFYNEEKDKKFQAVIETDDSD----EGGDEYCY----DGDDEDLLK 665  
 Db 2576 LEKAQQLQGLDEAKNNTIVLQSVKGLIQEVEDGKQKLEKKDDEETISRLNQIQDQEQLVS 2635

QY 666 PLKJDETDRFLMTLFFENNKKMLKRLAENPEY----ENEKLTKLRLNTIMEYQTRTEES 718  
 Db 2636 KLSQVEGEH--QLWKEQNLELRNLTVLEQKQVLQSKNALSQDTLEVQLOSSYKLNEN 2692

QY 719 ARGIIIFTKTROSAYALSOWITENEKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKF 778  
 Db 2693 ---LELTKMDRMSFV----EKVNKMTAKETELQREMHEMAQTAELQEEELSGEKNR 2741

QY 779 RTGKINLL----IATTVAEEGLDIKECNIVRYGLVTNEIAMVOARGRA 823  
 Db 2742 LAGELQOLLEBIKSSKQDQLKELTLENSSELLKSLDCMHQVKEEGKVRREEIAEYQLRILHE 2801

Qy 824 RADESTYVLVAHSGSGVIEHETVNDREKMMYKAIHCVQNMKPE----EYAHKILE 875  
 Db 2802 AEKXHQALLDDTNKQYEVEIQT---YREKLTSKE-ECLSSQKLEIDLLKSSKEELNNSLK 2857

Qy 876 LQMOSIMEMKMKTK----RNIAKHYKNKNPNSLITFLCKNCVSLACSGEDIHVIEK---925  
 Db 2858 ATTQ-ILLEELKTKMDNLKVYNQLKENERAOGGMKLLIKSCKOLE--EEKEILQKELS 2913

Qy 926 ----MEHVNMTPEFKELEYTVRENKA----LQKKCADYQINGEJICK 963  
 Db 2914 QLQAAQEOKQTGTVMDTKVDELTTIEKELKETLJEKTKEADEYLDKYCS-LLISHEKLEK 2972

Qy 964 CGQAWGTMWVH 974  
 Db 2973 AKEMLEJETQVAH 2983

RESULT 8  
 PCT-US95-16216-1  
 ; Sequence 1, Application PC/TUSS9516216  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rattner, Jerome B.  
 ; TITLE OF INVENTION: Nucleic Acid Encoding a Transiently  
 ; EXPRESSED Kinetochore Protein, and Methods of Use  
 ; NUMBER OF SEQUENCES: 4  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Dann, Dorfman, Herrell and Skillman  
 ; STREET: 1601 Market Street Suite 720  
 ; CITY: Philadelphia  
 ; STATE: PA  
 ; COUNTRY: USA  
 ; ZIP: 19103-2307  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US95/16216  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/353,700  
 ; FILING DATE: 09-DEC-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Reed, Janet E.  
 ; REGISTRATION NUMBER: 36,252  
 ; TELEPHONE: (215) 563-4100  
 ; TELEFAX: (215) 563-4044  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 3248 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: not relevant  
 ; TOPOLOGY: not relevant  
 ; MOLECULE TYPE: protein  
 ; HYPOTHETICAL: NO  
 ; ANTI-SENSE: NO  
 ; PCT-US95-16216-1

Query Match 3.6%; Score 191.5; DB 5; Length 3248;  
 Best Local Similarity 19.2%; Pred. No. 1.8e-06; Indels 281; Gaps 47;  
 Matches 221; Conservative 184; Missmatches 465; Indels 281; Gaps 47;

Qy 21 RVKMYIQVEPVL----DYLTFPLAEVKEQIORTVATSG----NMQAVELLIS 64  
 Db 1917 KIEACIELEKIVGELKKENSDSLSEKLEYFSCDHQELLQRVETSEGLNSDLLEMADKSSRE 1976

Qy 65 TLEKGVWHL--GWNTREFVEALRRRTGSPLAARYMNPELTDLPPSPSFENAHEDEYLQLNLLQ 122  
 Db 2742 LAGELQOLLEBIKSSKQDQLKELTLENSSELLKSLDCMHQVKEEGKVRREEIAEYQLRILHE 2801



Db 506 ALLQHKNAYQRKAHEADKKRNLENDVNSLKQLEDDKKRNQNSQISTEKVNQLQRQLD 565 CURRENT FILING DATE: 2001-10-12  
 Qy 293 EENVAARASPEPELQLRPYOMEVAQPA-----LEGNNTIICLPLTGSKTRVAVYIA 343 PRIORITY APPLICATION NUMBER: 60/240,409  
 Db 566 ETNALLRTESDTAARIKTKTOAESSKQIQQLESNNRDLQDKNCL--LETAKLKLEKEFINL 623 PRIORITY FILING DATE: 2000-10-12  
 Number of SEQ ID NOS: 1143  
 SOFTWARE: PERL Program  
 SEQ ID NO 296  
 LENGTH: 1388  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 OTHER INFORMATION: Incyte ID No. 6673549 064987CD1  
 FEATURE:  
 NAME/KEY: misc feature  
 OTHER INFORMATION: US-09-976-594-296

Query Match 3.5%; Score 188.5; DB 4; Length 1388;  
 Best Local Similarity 18.2%; Pred. No. 8.3e-07;  
 Matches 204; Conservative 195; Missmatches 399; Indels 321; Gaps 57;

Db 344 KDHLDKKKASEPGKVKIVLNKVLIVEQLFRKEFOPFLKKWYRVIGLSDTQLKISFPEV 403  
 624 QSALESEERRDRTHGSEII-----NDLQG-----RICGILEED-----654

Db 404 VKSCDIIISTAQI----LENSLMLLENGEDAGVQLS-DFSLIII-----DECHH-----447  
 655 LKNGKILLAKVELEKRLQERFTDLEK-EKSNEIMEIDMTYQLKVIQQSLEQEEAEHKATKA 713

Qy 448 -TNKEAVYNNIMRHYLMQKLKKENKPVIPPLPQILGLTASPGVGGATKQAKAEEH 505  
 714 RLADKNKIYESI----EEAKSEAMKEMKKLE-----ERTLKQKVENL 753

Db 506 IL---KLCANLDAFTIKTKVENLDQKLKNQIOPCCKFAIAADATREDPFKEKLEIMTRIQ 562  
 754 LLEAEKRCSSLDC-----DLKQSQQKCNELLKQ---KDVLINE-VRNLTKIEQETQ 801

Qy 563 TYCQMSPMSPMSDFGTQPYEQWAIQMERKAAKGNRCKERVAEHLRKYNEAL-----QI 613  
 802 KRC-LTQNDLKMOTQVNNTLKMSEKQLKQENNHLMEMKMKMLEKQNAELRKERQDAGQM 859

Qy 614 NDTIRMDAYTHLETFYN-----EKEKDK-----KPAVIEDDSD-----646  
 860 KELQDQLEAEQYFSTLYKTQVRELKKECEKTKLGKELLQQKKQELQDERDSLAAQLEITL 919

Db 647 -----EGGDDEXCDGDED----EDDLKKPL----KLDDETDRLMLTFFENNKKML 687  
 920 TKADSEQLARSIAEEQYSGLEKEKIMKELIEMMARHKQELTEKDATIASL-EETNRTL 978

Qy 688 KRLAENPEYENEKLTKLRTNTIMEQYTRTEE---SARGILIFTKTRQAYALSQWITEN--741  
 979 TSDVANLANEREEINNKLKDQVEQLSRJKEEISAAA-----KAQFEKQLLTERTLK 1031

Qy 742 -----EKEAEGVKKAHLIGAGHSSEPKEMTQNEQKEVISKFRTGKINLL-----786  
 Db 1032 TQAVNKLAEI-MNRKEPVKRGNDTDVR-RKEKENRKHMLIKSEREKLTKQMIKYQKELN 1089

Qy 787 -IATTVAEGLDIKECNIVRYGLVTNEIAMQARGRARA---DESTYVLVAHSGSGV 840  
 Db 1090 EMOAQIAEE---SQUIRIELQMTLDSKDSQDIEQLRSQQLQALHTGJDSSI-----GSGP 1139

Qy 841 IEHETVNDFREKMM--YKAIAHCVQNMKPEPEYAHKILELQMSIMEMKRMKTKRNIAKHYKN 898  
 1140 GDAEADDGFEPESRLEGWLSSLPRVNNNTKKFGWVKVIVVSSKKLIFYDSEQDKE-----QS 1194

Qy 899 NPSLITFLCKNCVIALCSGEDIHVLEKMHVNMTPEFKL---YIVRENKALQKCCADY 954  
 1195 NPVMVLDDIKLFLHVRPVTQDVFY-----RADAKEIPRIFQILYANEGESKKEQEF 1244

Db 955 QIN--GE---IIICKCQGQAWGTMVH-----974  
 1245 PVEPVGEKSNYICHKGHEFIPTLYHFPNTCEACMKPLWHMFKPAPPALCRRCHIKCHKDH 1304

Qy 975 ----KGDDLPCLKIRNFVYVFKN----NSTKKQYKKWV 1004  
 Db 1305 MDKKEEIIIAIAPC-KVYVDISTAKNLLLANSTEEQ-QKWW 1341

RESULT 10  
 US-09-976-594-296  
 Sequence 296, Application US/09976594  
 ; Patent No. 6673549  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Furness, Michael  
 ; INVENTOR: Buchbinder, Jenny  
 ; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELLS TREATED WITH STEROIDS  
 ; FILE REFERENCE: PA-0041 US  
 ; CURRENT APPLICATION NUMBER: US/09/976,594  
 ; CURRENT FILING DATE: 2001-10-12  
 ; PRIORITY APPLICATION NUMBER: 60/240,409  
 ; PRIORITY FILING DATE: 2000-10-12  
 ; NUMBER OF SEQ ID NOS: 1143  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 296  
 ; LENGTH: 1388  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; OTHER INFORMATION: Incyte ID No. 6673549 064987CD1  
 ; US-09-976-594-296  
 ; Sequence 296, Application US/09976594  
 ; Patent No. 6673549  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Furness, Michael  
 ; INVENTOR: Buchbinder, Jenny  
 ; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELLS TREATED WITH STEROIDS  
 ; FILE REFERENCE: PA-0041 US  
 ; CURRENT APPLICATION NUMBER: US/09/976,594  
 ; CURRENT FILING DATE: 2001-10-12  
 ; PRIORITY APPLICATION NUMBER: 60/240,409  
 ; PRIORITY FILING DATE: 2000-10-12  
 ; NUMBER OF SEQ ID NOS: 1143  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 296  
 ; LENGTH: 1388  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; OTHER INFORMATION: Incyte ID No. 6673549 064987CD1  
 ; US-09-976-594-296  
 ; Sequence 296, Application US/09976594  
 ; Patent No. 6673549  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Furness, Michael  
 ; INVENTOR: Buchbinder, Jenny  
 ; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELLS TREATED WITH STEROIDS  
 ; FILE REFERENCE: PA-0041 US  
 ; CURRENT APPLICATION NUMBER: US/09/976,594  
 ; CURRENT FILING DATE: 2001-10-12  
 ; PRIORITY APPLICATION NUMBER: 60/240,409  
 ; PRIORITY FILING DATE: 2000-10-12  
 ; NUMBER OF SEQ ID NOS: 1143  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 296  
 ; LENGTH: 1388  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; OTHER INFORMATION: Incyte ID No. 6673549 064987CD1  
 ; US-09-976-594-296  
 ; Sequence 296, Application US/09976594  
 ; Patent No. 6673549  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Furness, Michael  
 ; INVENTOR: Buchbinder, Jenny  
 ; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELLS TREATED WITH STEROIDS  
 ; FILE REFERENCE: PA-0041 US  
 ; CURRENT APPLICATION NUMBER: US/09/976,594  
 ; CURRENT FILING DATE: 2001-10-12  
 ; PRIORITY APPLICATION NUMBER: 60/240,409  
 ; PRIORITY FILING DATE: 2000-10-12  
 ; NUMBER OF SEQ ID NOS: 1143  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 296  
 ; LENGTH: 1388  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; OTHER INFORMATION: Incyte ID No. 6673549 064987CD1  
 ; US-09-976-594-296  
 ; Sequence 296, Application US/09976594  
 ; Patent No. 6673549  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Furness, Michael  
 ; INVENTOR: Buchbinder, Jenny  
 ; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELLS TREATED WITH STEROIDS  
 ; FILE REFERENCE: PA-0041 US  
 ; CURRENT APPLICATION NUMBER: US/09/976,594  
 ; CURRENT FILING DATE: 2001-10-12  
 ; PRIORITY APPLICATION NUMBER: 60/240,409  
 ; PRIORITY FILING DATE: 2000-10-12  
 ; NUMBER OF SEQ ID NOS: 1143  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 296  
 ; LENGTH: 1388  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; OTHER INFORMATION: Incyte ID No. 6673549 064987CD1  
 ; US-09-976-594-296  
 ; Sequence 296, Application US/09976594  
 ; Patent No. 6673549  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Furness, Michael  
 ; INVENTOR: Buchbinder, Jenny  
 ; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELLS TREATED WITH STEROIDS  
 ; FILE REFERENCE: PA-0041 US  
 ; CURRENT APPLICATION NUMBER: US/09/976,594  
 ; CURRENT FILING DATE: 2001-10-12  
 ; PRIORITY APPLICATION NUMBER: 60/240,409  
 ; PRIORITY FILING DATE: 2000-10-12  
 ; NUMBER OF SEQ ID NOS: 1143  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 296  
 ; LENGTH: 1388  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; OTHER INFORMATION: Incyte ID No. 6673549 064987CD1  
 ; US-09-976-594-296  
 ; Sequence 296, Application US/09976594  
 ; Patent No. 6673549  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Furness, Michael  
 ; INVENTOR: Buchbinder, Jenny  
 ; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELLS TREATED WITH STEROIDS  
 ; FILE REFERENCE: PA-0041 US  
 ; CURRENT APPLICATION NUMBER: US/09/976,594  
 ; CURRENT FILING DATE: 2001-10-12  
 ; PRIORITY APPLICATION NUMBER: 60/240,409  
 ; PRIORITY FILING DATE: 2000-10-12  
 ; NUMBER OF SEQ ID NOS: 1143  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 296  
 ; LENGTH: 1388  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; OTHER INFORMATION: Incyte ID No. 6673549 064987CD1  
 ; US-09-976-594-296  
 ; Sequence 296, Application US/09976594  
 ; Patent No. 6673549  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Furness, Michael  
 ; INVENTOR: Buchbinder, Jenny  
 ; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELLS TREATED WITH STEROIDS  
 ; FILE REFERENCE: PA-0041 US  
 ; CURRENT APPLICATION NUMBER: US/09/976,594  
 ; CURRENT FILING DATE: 2001-10-12  
 ; PRIORITY APPLICATION NUMBER: 60/240,409  
 ; PRIORITY FILING DATE: 2000-10-12  
 ; NUMBER OF SEQ ID NOS: 1143  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 296  
 ; LENGTH: 1388  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; OTHER INFORMATION: Incyte ID No. 6673549 064987CD1  
 ; US-09-976-594-296  
 ; Sequence 296, Application US/09976594  
 ; Patent No. 6673549  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Furness, Michael  
 ; INVENTOR: Buchbinder, Jenny  
 ; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELLS TREATED WITH STEROIDS  
 ; FILE REFERENCE: PA-0041 US  
 ; CURRENT APPLICATION NUMBER: US/09/976,594  
 ; CURRENT FILING DATE: 2001-10-12  
 ; PRIORITY APPLICATION NUMBER: 60/240,409  
 ; PRIORITY FILING DATE: 2000-10-12  
 ; NUMBER OF SEQ ID NOS: 1143  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 296  
 ; LENGTH: 1388  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; OTHER INFORMATION: Incyte ID No. 6673549 064987CD1  
 ; US-09-976-594-296  
 ; Sequence 296, Application US/09976594  
 ; Patent No. 6673549  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Furness, Michael  
 ; INVENTOR: Buchbinder, Jenny  
 ; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELLS TREATED WITH STEROIDS  
 ; FILE REFERENCE: PA-0041 US  
 ; CURRENT APPLICATION NUMBER: US/09/976,594  
 ; CURRENT FILING DATE: 2001-10-12  
 ; PRIORITY APPLICATION NUMBER: 60/240,409  
 ; PRIORITY FILING DATE: 2000-10-12  
 ; NUMBER OF SEQ ID NOS: 1143  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 296  
 ; LENGTH: 1388  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; OTHER INFORMATION: Incyte ID No. 6673549 064987CD1  
 ; US-09-976-594-296  
 ; Sequence 296, Application US/09976594  
 ; Patent No. 6673549  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Furness, Michael  
 ; INVENTOR: Buchbinder, Jenny  
 ; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELLS TREATED WITH STEROIDS  
 ; FILE REFERENCE: PA-0041 US  
 ; CURRENT APPLICATION NUMBER: US/09/976,594  
 ; CURRENT FILING DATE: 2001-10-12  
 ; PRIORITY APPLICATION NUMBER: 60/240,409  
 ; PRIORITY FILING DATE: 2000-10-12  
 ; NUMBER OF SEQ ID NOS: 1143  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 296  
 ; LENGTH: 1388  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; OTHER INFORMATION: Incyte ID No. 6673549 064987CD1  
 ; US-09-976-594-296  
 ; Sequence 296, Application US/09976594  
 ; Patent No. 6673549  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Furness, Michael  
 ; INVENTOR: Buchbinder, Jenny  
 ; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELLS TREATED WITH STEROIDS  
 ; FILE REFERENCE: PA-0041 US  
 ; CURRENT APPLICATION NUMBER: US/09/976,594  
 ; CURRENT FILING DATE: 2001-10-12  
 ; PRIORITY APPLICATION NUMBER: 60/240,409  
 ; PRIORITY FILING DATE: 2000-10-12  
 ; NUMBER OF SEQ ID NOS: 1143  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 296  
 ; LENGTH: 1388  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; OTHER INFORMATION: Incyte ID No. 6673549 064987CD1  
 ; US-09-976-594-296  
 ; Sequence 296, Application US/09976594  
 ; Patent No. 6673549  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Furness, Michael  
 ; INVENTOR: Buchbinder, Jenny  
 ; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELLS TREATED WITH STEROIDS  
 ; FILE REFERENCE: PA-0041 US  
 ; CURRENT APPLICATION NUMBER: US/09/976,594  
 ; CURRENT FILING DATE: 2001-10-12  
 ; PRIORITY APPLICATION NUMBER: 60/240,409  
 ; PRIORITY FILING DATE: 2000-10-12  
 ; NUMBER OF SEQ ID NOS: 1143  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 296  
 ; LENGTH: 1388  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; OTHER INFORMATION: Incyte ID No. 6673549 064987CD1  
 ; US-09-976-594-296  
 ; Sequence 296, Application US/09976594  
 ; Patent No. 6673549  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Furness, Michael  
 ; INVENTOR: Buchbinder, Jenny  
 ; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELLS TREATED WITH STEROIDS  
 ; FILE REFERENCE: PA-0041 US  
 ; CURRENT APPLICATION NUMBER: US/09/976,594  
 ; CURRENT FILING DATE: 2001-10-12  
 ; PRIORITY APPLICATION NUMBER: 60/240,409  
 ; PRIORITY FILING DATE: 2000-10-12  
 ; NUMBER OF SEQ ID NOS: 1143  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 296  
 ; LENGTH: 1388  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; OTHER INFORMATION: Incyte ID No. 6673549 064987CD1  
 ; US-09-976-594-296  
 ; Sequence 296, Application US/09976594  
 ; Patent No. 6673549  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Furness, Michael  
 ; INVENTOR: Buchbinder, Jenny  
 ; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELLS TREATED WITH STEROIDS  
 ; FILE REFERENCE: PA-0041 US  
 ; CURRENT APPLICATION NUMBER: US/09/976,594  
 ; CURRENT FILING DATE: 2001-10-12  
 ; PRIORITY APPLICATION NUMBER: 60/240,409  
 ; PRIORITY FILING DATE: 2000-10-12  
 ; NUMBER OF SEQ ID NOS: 1143  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 296  
 ; LENGTH: 1388  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; OTHER INFORMATION: Incyte ID No. 6673549 064987CD1  
 ; US-09-976-594-296  
 ; Sequence 296, Application US/09976594  
 ; Patent No. 6673549  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Furness, Michael  
 ; INVENTOR: Buchbinder, Jenny  
 ; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELLS TREATED WITH STEROIDS  
 ; FILE REFERENCE: PA-0041 US  
 ; CURRENT APPLICATION NUMBER: US/09/976,594  
 ; CURRENT FILING DATE: 2001-10-12  
 ; PRIORITY APPLICATION NUMBER: 60/240,409  
 ; PRIORITY FILING DATE: 2000-10-12  
 ; NUMBER OF SEQ ID NOS: 1143  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 296  
 ; LENGTH: 1388  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; OTHER INFORMATION: Incyte ID No. 6673549 064987CD1  
 ; US-09-976-594-296  
 ; Sequence 296, Application US/09976594  
 ; Patent No. 6673549  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Furness, Michael  
 ; INVENTOR: Buchbinder, Jenny  
 ; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELLS TREATED WITH STEROIDS  
 ; FILE REFERENCE: PA-0041 US  
 ; CURRENT APPLICATION NUMBER: US/09/976,594  
 ; CURRENT FILING DATE: 2001-10-12  
 ; PRIORITY APPLICATION NUMBER: 60/240,409  
 ; PRIORITY FILING DATE: 2000-10-12  
 ; NUMBER OF SEQ ID NOS: 1143  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 296  
 ; LENGTH: 1388  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; OTHER INFORMATION: Incyte ID No. 6673549 064987CD1  
 ; US-09-976-594-296  
 ; Sequence 296, Application US/09976594  
 ; Patent No. 6673549  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Furness, Michael  
 ; INVENTOR: Buchbinder, Jenny  
 ; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELLS TREATED WITH STEROIDS  
 ; FILE REFERENCE: PA-0041 US  
 ; CURRENT APPLICATION NUMBER: US/09/976,594  
 ; CURRENT FILING DATE: 2001-10-12  
 ; PRIORITY APPLICATION NUMBER: 60/240,409  
 ; PRIORITY FILING DATE: 2000-10-12  
 ; NUMBER OF SEQ ID NOS: 1143  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 296  
 ; LENGTH: 1388  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; OTHER INFORMATION: Incyte ID No. 6673549 064987CD1  
 ; US-09-976-594-296  
 ; Sequence 296, Application US/09976594  
 ; Patent No. 6673549  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Furness, Michael  
 ; INVENTOR: Buchbinder, Jenny  
 ; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELLS TREATED WITH STEROIDS  
 ; FILE REFERENCE: PA-0041 US  
 ; CURRENT APPLICATION NUMBER: US/09/976,594  
 ; CURRENT FILING DATE: 2001-10-12  
 ; PRIORITY APPLICATION NUMBER: 60/240,409  
 ; PRIORITY FILING DATE: 2000-10-12  
 ; NUMBER OF SEQ ID NOS: 1143  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 296  
 ; LENGTH: 1388  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; OTHER INFORMATION: Incyte ID No. 6673549 064987CD1  
 ; US-09-976-594-296  
 ; Sequence 296, Application US/09976594  
 ; Patent No. 6673549  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Furness, Michael  
 ; INVENTOR: Buchbinder, Jenny  
 ; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELLS TREATED WITH STEROIDS  
 ; FILE REFERENCE: PA-0041 US  
 ; CURRENT APPLICATION NUMBER: US/09/976,594  
 ; CURRENT FILING DATE: 2001-10-12  
 ; PRIORITY APPLICATION NUMBER: 60/240,409  
 ; PRIORITY FILING DATE: 2000-10-12  
 ; NUMBER OF SEQ ID NOS: 1143  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 296  
 ; LENGTH: 1388  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; OTHER INFORMATION:

QY 787 -IATTVAEEGLDIKECNIIVYGLVTNEIAMVQARGRARA-----DESTYVLVAAHSGSGV 840  
 ; :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 1090 EMQAQIAEE----SQRRIELQMTLDSKDS DIEQLRSQALHIGLDSSSI-----GSGP 1139

QY 841 IEHETVNDFREKOM--YKAITHCVQNMKPEEYAHKILELQMQSTIMEKMKTKRNIAKHYKN 898  
 ; :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 1140 GDAEADDGEPESRLEGWLSPVRNNTKKFGWVKVIVSSKKILFYDSEQDK-----QS 1194

QY 899 NPSLITFLCKNCSVLA CGEDIHVIEKOMHVNMTPEFKEL---YIVRENKALQOKKCADY 954  
 ; :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 1195 NPYMVLDDIKLFLHVVRPTQTDVY-----RADAKEIPRIFQILLYANEGESKEKEQEF 1244

QY 955 QIN--GE--I ICKCGQAWGTMW-----974  
 ; :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 1245 PVEPVGEKSNYI CHKGHEFIPTLYHFPTNCEACMKPLWEMFKPPPALECCRCH1KCHKD H 1304

QY 975 ---KGLDLPCCLKIRNFUVVFKN---NSTKPKQYKKWV 1004  
 ; :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 1305 MDKKEELIAPC-KVYDISTAKNLLANSTEEQ-QKWTW 1341

QY RESULT 11  
 US-09-919-172-29  
 ; Sequence 29, Application US/09919172  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Faris, Mary  
 ; INVENTION: PROSTATE CANCER MARKERS  
 ; FILE REFERENCE: PA-0036 US  
 ; CURRENT APPLICATION NUMBER: US/09/919,172  
 ; CURRENT FILING DATE: 2001-07-30  
 ; PRIOR APPLICATION NUMBER: 60/222,469  
 ; PRIOR FILING DATE: 2000-07-28  
 ; NUMBER OF SEQ ID NOS: 102  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 29  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; OTHER INFORMATION: Incyte ID No. 6673545 3774181CD1  
 ; LENGTH: 2125  
 ;  
 Query Match 3.5%; Score 187.5; DB 4; Length 2125;  
 Best Local Similarity 19.6%; Pred. No. 2e-06;  
 Matches 214; Conservative 164; Mismatches 375; Indels 337; Gaps 51;

QY 41 EVKEQIQRVTATSGNMQAVELLSTLEKGWHLGWTFEVLEARRT-----86  
 ; :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 468 ENOPENSKTLLATQLNQQ--KMLVSEIEMKQSKMDECQKYAEQYSATVQDYELQTMYRAM 525

QY 87 ---GSPLAARYMNP ETLDPSPSFENAHDEYLQQLNLLQPTLVDKLLVRDVLDKCMEE 141  
 ; :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 526 VDSQQKSPVKKRMRQSS ADLIIQE FMLRTRYTALVLTMTQYI--KFAGDSLKRLEEE 581

QY 142 ELLTIED-----RNRIAAENN--GNESGVRELLKRVQKENWFS AFLNVLR 186  
 ; :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 582 EIKRCKETSEHGAYSDLLQROKATVLENSKLTGKISELTERMVAELKKQSRVEELPKVR 641

QY 187 QTGNNEL-----VQELT----GSDCSESNAEIENL-SQV DGPQVE-EQLLSTTVQPN 232  
 ; :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 642 EAAENELRKQQRNVEDISLQKIRAESEA KQYRRELETIVREKEAARELLEVRQLTIEAE 701

QY 233 ---LEKEVWGMENNSSESSFA DSSV---VSESDTSLAEGSVSCLDESLGHNSNMGS DS 284  
 ; :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 702 AKRAAVEENLNFNRNQLEENTFTRRTLEDHILKRDLSL-----NDLQQQKNLMB E 752

QY 285 GTMGSDSDEENVAARASPEPELQLRPYQMEVAQPALEGK NII-----326  
 ; :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 753 IRRKRDNEELLKLIKQMEKDL---AFQKQVAEKQKIEKEQK KITEI QYTCREN 809

QY 327 ---ICLPTGSGKTRVAVYIARDHDKRKKA SEPGKVTIVL VN KVLLEQFLRKEFOPPLKK 383  
 ; :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 810 ALPVCPITOATSCRAVTGLOQEH--DKQKAEE----LKQQVDELTAA NRKAEOQDMREL 861

QY 384 WYR VIGLSGDTQLKISFPEVVKSCDII LISTAQILENSI LLENGEDAGVQLSDFSLI IID 443  
 ; :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 862 TYELN A---OLEKTSSE-----BKARLLKDKL-----D 887

QY 444 ECHHTNKEAVYNNIMRHYLMQKLKNRKKENKPVIPPLPQIQLGLTASPGGGATKQAKAE 503  
 ; :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 888 ET----NNTLR---CLKLELERKDQAEKGYSQQLRELGRLNQTTG-----KAE 929

QY 504 EHILKLCANLDAFTIKTVK-----ENLDQLKNQI QEPCKKFIA DATED REDP FK-----551  
 ; :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 930 E----AMQEASDLKKIRKNYQLELESLNHEKGKLQREVDTRAHAVA EKNIQHILNSQ 983

QY 552 ---EKILLEIMTRIQTYCQMS PMSDFGTQPYEWAQI QMEK- KLA KGN-----594  
 ; :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 984 IHSFRDEKELE---RLQ-ICQRK--SDH LKEQFEKSHEQ LQNTIAKEKENNNDK IQR LNEE 1037

QY 595 -RKERVCAEHL-----RKYN-----EALQI-----613  
 ; :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 1038 LEKSNECAEMLKQKV EELTRQNNETKLMMORIQA ESENTVLEKQTIQQRCEALKI QADGF 1097

QY 614 NDTIRMIDAYTHLETFYNEEKDKKFAVIEEDDSDEGGD-----DEYCDGDED EDDLK KPLK 668  
 ; :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 1098 KDQLRSTNEHLHKQT KTEQDFQRKIKC LEEDLA KSQNLVSEFKQCD-----QONII 1149

QY 669 LDETDRFLMTL FPE NN--KMLKRLAENP-----EYENELKLT KLRN-----TIMEQYT 713  
 ; :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 1150 I QNTKKEV RNLNA EK KRRGEQKVQLOQQA VQELNNR LKKVQ D EHLK TIEQMT 1209

QY 714 R-----TEESARGI LIFTKTR QAYA L SOWITEN EKFAEVG VKA HHLIGA HSSE FKPMT 767  
 ; :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 1210 HRYMVLQFEE SG-----KFK QSA-----EEFRK--KMEKLM-----ESKVIT 1244  
 QY 768 QNEQKEV I SKP RTG KIN LIA TTA E EGG D L I K E C N V I T R Y G L V T N E I A M V Q A R G R A D E 827  
 ; :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 1245 END-----ISGIRLDFVSLQ QENSRAQENAKLCETNI-----KELER QLQQYREQMQQ 1292

QY 828 STYVLVAH-SGSGVIEHETVNDREKMMYKA IHC VQNMKP-----EEYAHK LLELQMQS 880  
 ; :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 1293 GQHMEANHYQK CQK L E D E L I A Q K R E -----VENLKQMDQ I K E H H Q L V L L Q C E I 1343

QY 881 IMEKMKTKRNIAKHYKHN PSLITFLCKNCSVLA CGEDIH VIEKMH HVNMTPEFKELYI 940  
 ; :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 1344 -----QKKS TAKDCT EK PDF-----EMTVKECQHS GEL SS RNTG H L HPT P R S P L R W 1390

QY 941 VRENKA LQK K 950  
 ; :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 1391 TQE P QBLEK 1400

RESULT 12  
 US-08-973-462-B  
 ; Sequence 8, Application US/08973462B  
 ; Patent No. 6191270  
 ; GENERAL INFORMATION:  
 ; APPLICANT: DRUILHE, PIERRE  
 ; APPLICANT: DAUBERTIES, PIERRE  
 ; TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES  
 ; FILE REFERENCE: 0660-0125-0 PCT  
 ; CURRENT APPLICATION NUMBER: US/08/973, 462B  
 ; CURRENT FILING DATE: 1998-02-06  
 ; EARLIER APPLICATION NUMBER: PCT/FR96/00894  
 ; EARLIER FILING DATE: 1996-06-12  
 ; EARLIER APPLICATION NUMBER: FR 95/07007  
 ; EARLIER FILING DATE: 1995-06-13  
 ; NUMBER OF SEQ ID NOS: 29  
 ; SOFTWARE: PatentIn Ver.. 2.0  
 ; SEQ ID NO 8  
 ; LENGTH: 1786

TYPE: PRT	ORGANISM: Artificial Sequence	FEATURE: OTHER INFORMATION: Description of Artificial Sequence:Polypeptide	US-08-973-462-8
Query Match 3.5%; Score 186.5; DB 3; Length 1786;			
Best Local Similarity 19.6%; Pred. No. 1.9e-06;			
Matches 201; Conservative 152; Missmatches 373; Indels 301; Gaps 44;			
4.2 VKEQIORTVATSGNMQAVELLSTLEKGVWHLGWTREFVEALRRTGSPLAARYMNPELTD 101	QY	900 PSLITFLCKNCVLAACSGEDI -HVIKMHVNM -PEFKELYVRENKALLQKCADYQI 956	
4.99 VEEVAENVEESVAENVEEVAPTVEEIV-----APTVEEVAPSVE 541	Db	1265 KELEKALSEDSKEIIDAOKDTLEKVIEEHDTTLEVELKVEEDKI -EVSDLKD 1322	
RESULT 13	QY	957 NGEIICK 963	
US-09-150-867-1	Db	1323 LEEDILK 1329	
Sequence 1, Application US/09150867			
PATENT NO. 6645749			
GENERAL INFORMATION:			
APPLICANT: Wood, Kenneth W.			
APPLICANT: Sakowicz, Roman			
APPLICANT: Goldstein, Lawrence S.B.			
APPLICANT: Cleveland, Don W.			
APPLICANT: The Regents of the University of California			
TITLE OF INVENTION: Plus End-Directed Microtubule Motor Required for			
TITLE OF INVENTION: Chromosome Congression			
FILE REFERENCE: 185575C-000110US			
CURRENT APPLICATION NUMBER: US/09/150,867			
CURRENT FILING DATE: 1998-09-10			
EARLIER APPLICATION NUMBER: US 60/058,645			
EARLIER FILING DATE: 1997-09-11			
NUMBER OF SEQ ID NOS: 11			
SOFTWARE: Patentin Ver. 2.0			
SEQ ID NO 1			
LENGTH: 2954			
TYPE: PRT			
ORGANISM: Xenopus sp.			
FEATURE:			
OTHER INFORMATION: Xenopus centromere-associated protein-E (XCENP-E)			
OTHER INFORMATION: member of the kinesin superfamily of microtubule			
OTHER INFORMATION: motor proteins			
NAME/KEY: DOMAIN			
LOCATION: (1)..(472)			
OTHER INFORMATION: kinesin like motor domain			
FEATURE:			
NAME/KEY: DOMAIN			
LOCATION: (473)..(2752)			
OTHER INFORMATION: rod domain			
FEATURE:			
NAME/KEY: DOMAIN			
LOCATION: (2753)..(2954)			
OTHER INFORMATION: tail domain			
US-09-150-867-1			
Query Match 3.4%; Score 182.5; DB 4; Length 2954;	QY	6 STDENFRYLISCFPRARVKMYIQLVEPVLDYLTFLPAEVKEQIQRTVATSGNMQAVELLST 65	
Best Local Similarity 17.9%; Pred. No. 9.2e-06;	Db	564 SKEEEKTELVQSFELKI -----AELEEQL-----SVKAKNLEMVNTS 600	
Matches 182; Conservative 187; Mismatches 366; Indels 283; Gaps 42;			
Query Match 3.4%; Score 182.5; DB 4; Length 2954;	QY	66 LEKGWHLGWTREFVEALRRTGSPSRE----NAHDEYLQLINL 120	
Best Local Similarity 17.9%; Pred. No. 9.2e-06;	Db	601 REHSINAEVQTDVEKEVVRKEMSVLGDSGYNAASNDSLQDSSVGDKRLSSSHDECIEHRM 660	
Matches 182; Conservative 187; Mismatches 366; Indels 283; Gaps 42;			
Query Match 3.4%; Score 182.5; DB 4; Length 2954;	QY	121 LQPTLV-----KLLVRDVLDKCMEEELLTIDRN--RRIAAANN-- 158	
Best Local Similarity 17.9%; Pred. No. 9.2e-06;	Db	661 LEQKIVDLEEFIENLNKKSENDKQKSSEQDFMESISQLCATMAEKANALEELALMRDNFD 720	
Matches 182; Conservative 187; Mismatches 366; Indels 283; Gaps 42;			
Query Match 3.4%; Score 182.5; DB 4; Length 2954;	QY	159 ----GNEGVREL--LKRIV-- -QKENWESAFNVLROTGNNELVQELTGSDCSESNAEI 209	
Best Local Similarity 17.9%; Pred. No. 9.2e-06;	Db	721 NITLENETLKREIADLERSLKENQETNEPEILLEKETQKEHEAQLTHEIGSLKKLVNAEM 780	
Matches 182; Conservative 187; Mismatches 366; Indels 283; Gaps 42;			
Query Match 3.4%; Score 182.5; DB 4; Length 2954;	QY	210 ENLSQVDGPQVEEQLL-STTVQ-----WGMENNSSESSFFADS 252	
Best Local Similarity 17.9%; Pred. No. 9.2e-06;	Db		
Matches 182; Conservative 187; Mismatches 366; Indels 283; Gaps 42;			

Query Match 3.4%; Score 179; DB 4; Length 666;  
 Best Local Similarity 19.5%; Pred. No. 1.7e-06;  
 Matches 137; Conservative 114; Mismatches 230; Indels 220; Gaps 35;

Y 253 SVVSEDTSLAEGSVSC-----LDESLGHNSNMGS-GTMS-----DSDEENV- 296  
 Y 841 QSLSDAEAVTRDAQKBCSFLRSENLEKEKMEDTSNWYNOKEKAASLFEKOLETEKSNYK 900  
 Y 297 AARASPEPELQLRPyQMEAQPALEGKNIICLPTGSKTRAVVYIAKDHLK--KKKA 353  
 Y 901 KMEADLQKELQSAFNEINYNLLAGK-----VPRDILSRVELEKKV 942  
 Y 354 SEPGBKVI-----VLVNVKULLVEQLFRKEFQPFLLKWKYRVLGSGDTQLKISFPEVVK 405  
 Y 943 SEFSKQLEKALEEKNALENEVTCLESEY---KFLP-----NEVECLKNQISK 985  
 Y 406 SCDDIIISTAQILLENSLLINLENGEDAGGVQLSDFSLIIIDECHTNKEAVNNIMRHYLMQK 465  
 Y 986 ASEEEIMLLKQEGEHSAASISIKSQEIIIMQEQSEQLQLTDEVHTQSQ-VQOTEEQYLEMKK 1044  
 Y 466 LKNNRLEKK--ENKPVIPPLPQILGLTASPGVGGATKQAKAAEHILKLCANLDAFTIKTVKE 523  
 Y 524 NLQDQLKNOIQCPCRKKAFAADATREDFPEKXLE---IMTRIQTYCQMSPMSDFGTQPY 579  
 Y 1080 KIADTKHELEE---TIRD---SEAED-LLREMENLKG-TMSEVY 1079  
 Y 580 QWAQIMEKKAQKGNRKEVRVCAEHLRKYNEALQINDTIRMDAYTH---LETFYNE 632  
 Y 1121 ---SLPPSKLVEGNSQ-----DPIEINDYHNLLALATERNNIMUVCLETERNS 1164  
 Y 633 EKDK-----KFAVIEEDDSDEGGDDEYCDGDEDDEDDLKKPLKLDDETDRFLMTLEFFNNQM 686  
 Y 1165 LKEQVIDLNTQLQSLQAQSIEKSQDQPKDQLEGEVKLILLEME---LLKGHLTDSQLS 1220  
 Y 687 LKRLAENPEYENBKLTKLNRNTIMEQYTR---TEESARGILIFTKTROSAYALSQWITENEKF 744  
 Y 1221 IEKL---QLENLEVTEKQLQTQEMKNNTIERNEQTNFEDLKAEDSLKQDLSBN--- 1273  
 Y 745 AEVGVKAHHLIGAGHSSEFKPMTNQKEVISKFR----- 779  
 Y 1274 IEQSIETQDELRAAQEE---LREQQLVDSFRQQLLDCSGVGISSPNHDAVANQEKVS 1327  
 Y 780 TGKINNLIATTVAEEGLDIK-ECNIVIRYGLVNEIAMVQARGRADAESTYVLAHSGS 83B  
 Y 1328 LGEVNSLQSEMRLGERDELQTSCKALV-----SELELLRAHKVSVEGENLEITKLLNG- 1380  
 Y 839 GYIEHETVNDFREKMMYKAHCVQNMK-----PEEYAHKILELQMQSIMEKMK 887  
 Y 1381 -LEKEIILGKSEEESEVLSKM--LENLKDENNKLKEQAEYSSKENQFSLEEVSFGSQK 1434  
 Y 601 DVISATVESDETNOQQQTTELPKKMTKKEQKTIENIEKEMK 641

RESULT 14  
 US-09-134-001C-5465  
 Sequence 5465, Application US/09134001C  
 Patent No. 6380370  
 GENERAL INFORMATION:  
 APPLICANT: Lynn Doucette-Stamm et al  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
 FILE REFERENCE: GTC-007  
 CURRENT APPLICATION NUMBER: US/09/134,001C  
 CURRENT FILING DATE: 1998-08-13  
 PRIOR APPLICATION NUMBER: US 60/064,964  
 PRIOR FILING DATE: 1997-11-08  
 PRIOR APPLICATION NUMBER: US 60/055,779  
 PRIOR FILING DATE: 1997-08-14  
 NUMBER OF SEQ ID NOS: 5674  
 SEQ ID NO: 5465  
 LENGTH: 666  
 TYPE: PRT  
 ORGANISM: *Staphylococcus epidermidis*  
 US-09-134-001C-5465

RESULT 15  
 US-08-685-576-1  
 Sequence 1, Application US/08685576  
 Patent No. 5906819  
 GENERAL INFORMATION:  
 APPLICANT: Kaibuchi, Kozo  
 APPLICANT: Iwamatsu, Akihiro  
 APPLICANT: Nakano, Takeshi  
 APPLICANT: Ito, Masaaki  
 APPLICANT: Takahashi, No. 5906819uak1  
 TITLE OF INVENTION: RHO TARGET PROTEIN RHO-KINASE  
 NUMBER OF SEQUENCES: 16  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Foley & Lardner  
 STREET: 3000 K Street, N.W., Suite 500  
 CITY: Washington  
 STATE: D.C.  
 COUNTRY: USA  
 ZIP: 20007-5109  
 COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/685,576  
; FILING DATE: 24-JUL-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 7-325129  
; FILING DATE: 20-NOV-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 8-17150  
; FILING DATE: 05-JAN-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 8-131206  
; FILING DATE: 26-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bent, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 16887/843  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 672-5300  
; TELEFAX: (202) 672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1388 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-685-576-1

Query Match 3.3%; Score 174; DB 2; Length 1388;  
 Best Local Similarity 18.5%; Pred. No. 1.5e-05;  
 Matches 207; Conservative 186; Mismatches 390; Indels 334; Gaps 58;

QY 83 LRRTGSPILLAARYMNPEL-TDLPPSPSFENAHDYLQQLNLLQPTLV-----D 127  
 Db 364 IRETAAPVY----PELSSDIDSNFDDIEDDKGDVETFPIPKAFVGNQLPFIGFTYRE 418

QY 128 KLLVRDVLDKCMEEELLTIEDNRRIAANNGNEGVRELLKRVQENWFAFLNVLRQ 187  
 Db 419 NLLISD-SPSCKEND--SIQSRAK----NEES-QEIQKLYTLEELHSLTEIQ----461

QY 188 TGNNELYQELTGSDCSESNAEENLSQLVDGQPOVEEQLLSTTVQPNLKEVWGMENSS 247  
 Db 462 -AKEELEQK----CKSVNTRLEKVAK---ELEEEI---TIRKNESTLQLEREKALL 508

QY 248 SFADSSVYSESSTDLSL----AEGSVSCLDSL----GHNSNMGSDS-GTMGSDSDEEN 295  
 Db 509 QHKNAEYQRKADHEADKRNLENDVNSLKDQLEDLKRNQNSQISTEKVNQQLQRQDETN 568

QY 296 VAARASPEPELQRLRQYOMEVAQPA-----LEGKNIITICLPTGSGKTRVAVYIAKDH 346  
 Db 569 ALLRTESDTAAIRKTKQAESSKQIQQLESNNRDLQDN---CL---LETAKLK 615

QY 347 LDKKKKASEPGKVTVLYKVLVEQLFRKEFOPFLKKWRYRVLGSGDTQLKIS-FPEVVK 405  
 Db 616 LEKE-----FINLQSVLESERRDRTHGSEII-----NDLQGRISGLEEDVK 656

QY 406 SCDDIIISTAQI----LENSLLNLENGEDAGVQLS-DFSLIII----DECHH----447  
 Db 657 NGKILLAKVELEKRLQQLQERFTDLEK-EKNNMEIDMTYQLKVIQQSLEQEETEHKATKARL 715

QY 448 TNKEAVYNNIMRHYLMQKLKNRLLKBNKPVIPLPQILGLTASPPVGGATKQAKAEEHIL 507  
 Db 716 ADKNKIYESI-----EAKSEAMKEMEKLSE-----ERTLKQKVENLL 755

QY 508 ---KLCANLDAFTIKTKVENLDDQLKNQIOPCKKFAIAADATREDDPFKEKLLIMTRIOTY 564  
 Db 756 EAERKRCSSILD----DLKQSQOKNELLQ----KDVLINEQETOKR 803

GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2004, 19:11:53 ; Search time 54 Seconds  
 (without alignments)  
 5340.225 Million cell updates/sec

Title: US-09-515-363C-2  
 Perfect score: 5311 MSNGYSTDENFRYLISCFRA.....LPITFPNLDYSECCLFSDDED 1025

Sequence: Scoring table: BLOSUM62  
 Gapext 0.5

Searched: 1155919 seqs, 281338677 residues  
 Total number of hits satisfying chosen parameters: 1155919

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database :

1: /cgn2\_6/ptodata/2/pubpaa/'US07 PUBCOMB.pep:\*

2: /cgn2\_6/ptodata/2/pubpaa/'PCT NEW PUB.pep:\*

3: /cgn2\_6/ptodata/2/pubpaa/'US06 NEW PUB.pep:\*

4: /cgn2\_6/ptodata/2/pubpaa/'US06 PUBCOMB.pep:\*

5: /cgn2\_6/ptodata/2/pubpaa/'PCTUS PUBCOMB.pep:\*

6: /cgn2\_6/ptodata/2/pubpaa/'PCT NEW PUB.pep:\*

7: /cgn2\_6/ptodata/2/pubpaa/'US08 NEW PUB.pep:\*

8: /cgn2\_6/ptodata/2/pubpaa/'US07 NEW PUB.pep:\*

9: /cgn2\_6/ptodata/2/pubpaa/'US09A PUBCOMB.pep:\*

10: /cgn2\_6/ptodata/2/pubpaa/'US09B PUBCOMB.pep:\*

11: /cgn2\_6/ptodata/2/pubpaa/'US09C PUBCOMB.pep:\*

12: /cgn2\_6/ptodata/2/pubpaa/'US09 NEW PUB.pep:\*

13: /cgn2\_6/ptodata/2/pubpaa/'US10A PUBCOMB.pep:\*

14: /cgn2\_6/ptodata/2/pubpaa/'US10B PUBCOMB.pep:\*

15: /cgn2\_6/ptodata/2/pubpaa/'US10C PUBCOMB.pep:\*

16: /cgn2\_6/ptodata/2/pubpaa/'US10 NEW PUB.pep:\*

17: /cgn2\_6/ptodata/2/pubpaa/'US60 NEW PUB.pep:\*

18: /cgn2\_6/ptodata/2/pubpaa/'US60 PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5311	100.0	1025	14	US-10-055-475-2
2	5311	100.0	1025	14	US-10-228-897-2
3	5307	99.9	1025	14	US-10-055-475-9
4	5299	99.8	1025	14	US-10-055-475-7
5	5285	99.5	1025	16	US-10-275-822A-2
6	2456	46.2	468	15	US-10-094-749-2467
7	1087.5	20.5	925	16	US-10-408-765A-2031
8	1076	20.3	486	14	US-10-055-475-12
9	1040.5	19.6	487	15	US-10-108-260A-4515
10	788.5	14.8	514	14	US-10-055-475-13
11	783	14.7	166	9	US-09-864-761-45372
12	782	14.7	154	14	US-10-106-698-5088
13	782	14.7	154	15	US-10-264-049-2244
14	617	11.6	416	14	US-10-055-475-14
15	563.5	10.6	447	15	US-10-264-049-2303

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5311	100.0	1025	14	US-10-055-475-2
2	5311	100.0	1025	14	US-10-228-897-2
3	5307	99.9	1025	14	US-10-055-475-9
4	5299	99.8	1025	14	US-10-055-475-7
5	5285	99.5	1025	16	US-10-275-822A-2
6	2456	46.2	468	15	US-10-094-749-2467
7	1087.5	20.5	925	16	US-10-408-765A-2031
8	1076	20.3	486	14	US-10-055-475-12
9	1040.5	19.6	487	15	US-10-108-260A-4515
10	788.5	14.8	514	14	US-10-055-475-13
11	783	14.7	166	9	US-09-864-761-45372
12	782	14.7	154	14	US-10-106-698-5088
13	782	14.7	154	15	US-10-264-049-2244
14	617	11.6	416	14	US-10-055-475-14
15	563.5	10.6	447	15	US-10-264-049-2303

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB ID	Description
1	5311	100.0	1025	14	US-10-055-475-2
2	5311	100.0	1025	14	US-10-228-897-2
3	5307	99.9	1025	14	US-10-055-475-9
4	5299	99.8	1025	14	US-10-055-475-7
5	5285	99.5	1025	16	US-10-275-822A-2
6	2456	46.2	468	15	US-10-094-749-2467
7	1087.5	20.5	925	16	US-10-408-765A-2031
8	1076	20.3	486	14	US-10-055-475-12
9	1040.5	19.6	487	15	US-10-108-260A-4515
10	788.5	14.8	514	14	US-10-055-475-13
11	783	14.7	166	9	US-09-864-761-45372
12	782	14.7	154	14	US-10-106-698-5088
13	782	14.7	154	15	US-10-264-049-2244
14	617	11.6	416	14	US-10-055-475-14
15	563.5	10.6	447	15	US-10-264-049-2303

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB ID	Description
1	5311	100.0	1025	14	US-10-055-475-2
2	5311	100.0	1025	14	US-10-228-897-2
3	5307	99.9	1025	14	US-10-055-475-9
4	5299	99.8	1025	14	US-10-055-475-7
5	5285	99.5	1025	16	US-10-275-822A-2
6	2456	46.2	468	15	US-10-094-749-2467
7	1087.5	20.5	925	16	US-10-408-765A-2031
8	1076	20.3	486	14	US-10-055-475-12
9	1040.5	19.6	487	15	US-10-108-260A-4515
10	788.5	14.8	514	14	US-10-055-475-13
11	783	14.7	166	9	US-09-864-761-45372
12	782	14.7	154	14	US-10-106-698-5088
13	782	14.7	154	15	US-10-264-049-2244
14	617	11.6	416	14	US-10-055-475-14
15	563.5	10.6	447	15	US-10-264-049-2303

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB ID	Description
1	5311	100.0	1025	14	US-10-055-475-2
2	5311	100.0	1025	14	US-10-228-897-2
3	5307	99.9	1025	14	US-10-055-475-9
4	5299	99.8</			

RESULT 2  
US-10-228-897-2  
; Sequence 2, APP  
; Publication No.  
; GENERAL INFORMATION  
; APPLICANT: Fis  
; APPLICANT: Kae  
; APPLICANT: Goc  
; TITLE OF INVEN

RESULT 3  
US-10-055-475-9  
; Sequence 9, Application US/10055475  
; Publication No. US20030022855A1  
; GENERAL INFORMATION:  
; APPLICANT: Fisher, Paul B.  
; APPLICANT: Kang, Dong-Chul  
; APPLICANT: Gopalkrishnan, Rahul V.  
; TITLE OF INVENTION: USE OF MDA-5 AS AN ANTIVIRAL AND  
; TITLE OF INVENTION: ANTIPROLIFERATIVE AGENT  
; FILE REFERENCE: A34614-A-PCT-USA-A (070050-1921)  
; CURRENT APPLICATION NUMBER: US/10/055, 475  
; CURRENT FILING DATE: 2002-01-22  
; PRIOR APPLICATION NUMBER: PCT/US01/06960  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: 09/515, 363  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 9  
; LENGTH: 1025  
; TYPE: PRT  
; ORGANISM: homo sapiens  
; US-10-055-475-9

Query Match 99.9%; Score 5307; DB 14; Length 1025;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1024; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

RESULT 4  
US-10-055-475-7  
; Sequence 7, Application US/10055475  
; Publication No. US20030022855A1  
; GENERAL INFORMATION:  
; APPLICANT: Fisher, Paul B.  
; APPLICANT: Kang, Dong-Chul  
; APPLICANT: Gopalkrishnan, Rahul V.  
; TITLE OF INVENTION: USE OF MDA-5 AS AN ANTIVIRAL AND  
; TITLE OF INVENTION: ANTIPROLIFERATIVE AGENT  
; FILE REFERENCE: A34614-A-PCT-USA-A (070050-1921)  
; CURRENT APPLICATION NUMBER: US/10/055, 475  
; CURRENT FILING DATE: 2002-01-22  
; PRIOR APPLICATION NUMBER: PCT/US01/06960  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: 09/515, 363  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 1025  
; TYPE: PRT  
; ORGANISM: homo sapiens  
; US-10-055-475-7

Query Match 99.9%; Score 5307; DB 14; Length 1025;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1024; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Query 1 MSNGYSTDENFRYLISCFRARMYIQLVEPVLDYLTFPLPAEVKEQIQRSTVATSGNMQAVE 60  
Db 1 MSNGYSTDENFRYLISCFRARMYIQLVEPVLDYLTFPLPAEVKEQIQRSTVATSGNMQAVE 60  
Query 61 LLLSTLEKGVWHLGWTRFVEALRTGSPLAARYMNPPELTDLPSSENAHDEYLQLLNL 120  
Db 61 LLLSTLEKGVWHLGWTRFVEALRTGSPLAARYMNPPELTDLPSSENAHDEYLQLLNL 120  
Query 121 LQPTLYDKLLYRVDVLDKCMEEELLTIEDRNRRIAAENNGNEGVRELLKRIYQKENWFA 180  
Db 121 LQPTLYDKLLYRVDVLDKCMEEELLTIEDRNRRIAAENNGNEGVRELLKRIYQKENWFA 180  
Query 181 FLNVLROTGNNELVQELTGSDCSESNAEIEINLSQVDGPQVEEQLLTTVQPNLEKEVWGM 240  
Db 181 FLNVLROTGNNELVQELTGSDCSESNAEIEINLSQVDGPQVEEQLLTTVQPNLEKEVWGM 240  
Query 241 ENNSSESSFADSSVVSESDTSLAEGSVSCLDESIGHNSNMGSDSGTMGSDSDEENVAARA 300  
Db 241 ENNSSESSFADSSVVSESDTSLAEGSVSCLDESIGHNSNMGSDSGTMGSDSDEENVAARA 300  
Query 301 SPEPELQJRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDHLDKKKKASEPGKVI 360  
Db 301 SPEPELQJRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDHLDKKKKASEPGKVI 360  
Query 541 IADATREDPFKEKLLEIMTRIQTYCCQMSPMSPMSDFGTQPYEQWAIQMEKKAEGNRKERV 600  
Db 541 IADATREDPFKEKLLEIMTRIQTYCCQMSPMSPMSDFGTQPYEQWAIQMEKKAEGNRKERV 600  
Query 601 AEHLRKYNEALQINDTIRIMDAYTHLETFYNEEKDKKFAVIEDDSDEGGDDEYCDGDED 660  
Db 601 AEHLRKYNEALQINDTIRIMDAYTHLETFYNEEKDKKFAVIEDDSDEGGDDEYCDGDED 660  
Query 661 DLLKKPLKLDETDRLMTLFFENNKMLKRLAENPEYENEKLTKLRTNTIMEQYTRTTEESAR 720  
Db 661 DLLKKPLKLDETDRLMTLFFENNKMLKRLAENPEYENEKLTKLRTNTIMEQYTRTTEESAR 720  
Query 721 GIIFTKTRQSAVALSOWITENEKFAEVGYKAHHLIGAHSSEFKPMTONEQKEVISKFR 780  
Db 721 GIIFTKTRQSAVALSOWITENEKFAEVGYKAHHLIGAHSSEFKPMTONEQKEVISKFR 780  
Query 781 GKIINLIATTVAEGLDIKECNIVIRGLYTNEIAMVQARGRADESTYVLYAHSGSGV 840  
Db 781 GKIINLIATTVAEGLDIKECNIVIRGLYTNEIAMVQARGRADESTYVLYAHSGSGV 840  
Query 841 IEHETVNDFREKMYKAITHCQVNMKPPEEYAHKILEQVKAHHLIGAHSSEFKPMTONEQKEVISKFR 900  
Db 841 IEHETVNDFREKMYKAITHCQVNMKPPEEYAHKILEQVKAHHLIGAHSSEFKPMTONEQKEVISKFR 900  
Query 901 SLITFLCKNCNSVLAQSGEDIHVIEKMHVNMTPPEFKELYIVRENKAQKKCDAFTYVLYAHSGSGV 960  
Db 901 SLITFLCKNCNSVLAQSGEDIHVIEKMHVNMTPPEFKELYIVRENKAQKKCDAFTYVLYAHSGSGV 960  
Query 961 ICKCGQAWGTMVHKGLDPLCKIRNFYVVFKNNSTKQYKKWELPITFPNLIDYSECC 1020  
Db 961 ICKCGQAWGTMVHKGLDPLCKIRNFYVVFKNNSTKQYKKWELPITFPNLIDYSECC 1020

Query Match 99.8%; Score 5299; DB 14; Length 1025;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1023; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSNGYSTDENFRYLISCFRARVMYIQLVEPVLDYLTLPAEVKEQIORTVATSGNMQAVE 60  
Db 1 MSNGYSTDENFRYLISCFRARVMYIQLVEPVLDYLTLPAEVKEQIORTVATSGNMQAVE 60

Qy 61 LLLSTLEKGWHLGWTREFVEALRRRTGSPLAARYMNPELTDLPSFSFENAHDYELQOLLNL 120  
Db 61 LLLSTLEKGWHLGWTREFVEALRRRTGSPLAARYMNPELTDLPSFSFENAHDYELQOLLNL 120

Qy 121 LQPTLVDKLRLVRDVLDKCMEEELLTIEDNRRTAAEENNNGESGVRELLKRIVOKENWFA 180  
Db 121 LQPTLVDKLRLVRDVLDKCMEEELLTIEDNRRTAAEENNNGESGVRELLKRIVOKENWFA 180

Qy 181 FLNVLRQTGNNELVQELTGSDCSESNAEIEINLSQLDGPQVEEQLLSTTQPNLKEVWGM 240  
Db 181 FLNVLRQTGNNELVQELTGSDCSESNAEIEINLSQLDGPQVEEQLLSTTQPNLKEVWGM 240

Qy 241 ENNSSESSFADESSVVYSESSTDTSIAEGSVSCLDESIGHNSNMGSDSGTMGSDSDEENVAARA 300  
Db 241 ENNSSESSFADESSVVYSESSTDTSIAEGSVSCLDESIGHNSNMGSDSGTMGSDSDEENVAARA 300

Qy 301 SPEPELQLRPYQMEVAQPALEGKNIITICLPTGSGKTRVAVYIAKDHLDKKKKASEPGKVI 360  
Db 301 SPEPELQLRPYQMEVAQPALEGKNIITICLPTGSGKTRVAVYIAKDHLDKKKKASEPGKVI 360

Qy 361 VLVNKVLLVEQLFRKEFQPFLLKWKWVIGLGSQDTQLKISFPEVVKSCDIIISTAQILENS 420  
Db 361 VLVNKVLLVEQLFRKEFQPFLLKWKWVIGLGSQDTQLKISFPEVVKSCDIIISTAQILENS 420

Qy 421 LNNLENGEDAGVQLSDLSLIIIDECHHTNKEAVYNNIMRHYLMOKLKNRNLKKENKPVIP 480  
Db 421 LNNLENGEDAGVQLSDLSLIIIDECHHTNKEAVYNNIMRHYLMOKLKNRNLKKENKPVIP 480

Qy 481 LPQILGLTASPGVGGATKQAKAEEHILKLCANLDAFTIKTVKENLDQLKNOIQLEPCKKFA 540  
Db 481 LPQILGLTASPGVGGATKQAKAEEHILKLCANLDAFTIKTVKENLDQLKNOIQLEPCKKFA 540

Qy 541 IADATREDPFEKLLIMTRIQTYCQMSPMSDFGTQPEQWAIQMEKKAAKKGKGNRKERV 600  
Db 541 IADATREDPFEKLLIMTRIQTYCQMSPMSDFGTQPEQWAIQMEKKAAKKGKGNRKERV 600

Qy 601 AEHLRKYNEALQINDTIRMIDAYTHLETFYNEEKDKKFAVIEDDSDEGGDDEYCDGDEDE 660  
Db 601 AEHLRKYNEALQINDTIRMIDAYTHLETFYNEEKDKKFAVIEDDSDEGGDDEYCDGDEDE 660

Qy 661 DDLKKPKPLKDETDRFLMLTFENNPKMLKRLAENPEYENEKLTKLRLNTIMEQYTRTEESAR 720  
Db 661 DDLKKPKPLKDETDRFLMLTFENNPKMLKRLAENPEYENEKLTKLRLNTIMEQYTRTEESAR 720

Qy 721 GIIIFTKTRQSYALSQWITENEKFAEVGVKAHHLTGAGHSSEFKPMTONEQKEVISKFR 780  
Db 721 GIIIFTKTRQSYALSQWITENEKFAEVGVKAHHLTGAGHSSEFKPMTONEQKEVISKFR 780

Qy 781 GKNLIIATTVAEGLDIKECNIVIRYGLVTNEIAMVQARGRAADESTYVLVAHSGSGV 840  
Db 781 GKNLIIATTVAEGLDIKECNIVIRYGLVTNEIAMVQARGRAADESTYVLVAHSGSGV 840

Qy 841 JEHETYNDFREKMMYKAIIHCYQNMKPEEYAHKILELQMOQSIMEKKMKTKNIAKHYKNNP 900  
Db 841 IERETVNDFREKMMYKAIIHCYQNMKPEEYAHKILELQMOQSIMEKKMKTKNIAKHYKNNP 900

Qy 901 SLITFLCKNCSVLAACSGEDIHVTEKMHVNMTPEFKELYIVRENKALQKKCADYQINGEI 960  
Db 901 SLITFLCKNCSVLAACSGEDIHVTEKMHVNMTPEFKELYIVRENKALQKKCADYQINGEI 960

Qy 961 ICKCGQAWGTMVHKGLDLPCLKIRNFVVFKNINSTKQYKKWVELPITFPNLDSYSECC 1020  
Db 961 ICKCGQAWGTMVHKGLDLPCLKIRNFVVFKNINSTKQYKKWVELPITFPNLDSYSECC 1020

Qy 541 IADATREDPFEKLLIMTRIQTYCQMSPMSDFGTQPEQWAIQMEKKAAKKGKGNRKERV 600  
Db 541 IADATREDPFEKLLIMTRIQTYCQMSPMSDFGTQPEQWAIQMEKKAAKKGKGNRKESVC 600

RESULT 5  
US-10-275-822A-2  
; Sequence 2, Application US/10275822A  
; Publication No. US20040086500A1  
; GENERAL INFORMATION:  
; APPLICANT: Bahr, Georges  
; APPLICANT: Coudre, Cecile  
; APPLICANT: Capron, Andre  
; TITLE OF INVENTION: RH116 Polypeptide and its Fragments and Polynucleotides Encoding  
; its Fragments and Therapeutic Uses  
; FILE REFERENCE: 017753-170  
; CURRENT APPLICATION NUMBER: US/10/275,822A  
; CURRENT FILING DATE: 2002-11-12  
; PRIOR APPLICATION NUMBER: PCT/FR01/01441  
; PRIOR FILING DATE: 2001-05-11  
; PRIOR APPLICATION NUMBER: FR 00/06,030  
; PRIOR FILING DATE: 2000-05-11  
; NUMBER OF SEQ ID NOS: 56  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO: 2  
; LENGTH: 1025  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-275-822A-2

Query Match 99.5%; Score 5285; DB 16; Length 1025;  
Best Local Similarity 99.5%; Pred. No. 0;  
Matches 1020; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MSNGYSTDENFRYLISCFRARVMYIQLVEPVLDYLTLPAEVKEQIORTVATSGNMQAVE 60  
Db 1 MSNGYSTDENFRYLISCFRARVMYIQLVEPVLDYLTLPAEVKEQIORTVATSGNMQAVE 60

Qy 61 LLLSTLEKGWHLGWTREFVEALRRRTGSPLAARYMNPELTDLPSFSFENAHDYELQOLLNL 120  
Db 61 LLLSTLEKGWHLGWTREFVEALRRRTGSPLAARYMNPELTDLPSFSFENAHDYELQOLLNL 120

Qy 121 LQPTLVDKLRLVRDVLDKCMEEELLTIEDNRRTAAEENNNGESGVRELLKRIVOKENWFA 180  
Db 121 LQPTLVDKLRLVRDVLDKCMEEELLTIEDNRRTAAEENNNGESGVRELLKRIVOKENWFA 180

Qy 181 FLNVLRQTGNNELVQELTGSDCSESNAEIEINLSQLDGPQVEEQLLSTTQPNLKEVWGM 240  
Db 181 FLNVLRQTGNNELVQELTGSDCSESNAEIEINLSQLDGPQVEEQLLSTTQPNLKEVWGM 240

Qy 241 ENNSSESSFADESSVVYSESSTDTSIAEGSVSCLDESIGHNSNMGSDSGTMGSDSDEENVAARA 300  
Db 241 ENNSSESSFADESSVVYSESSTDTSIAEGSVSCLDESIGHNSNMGSDSGTMGSDSDEENVAARA 300

Qy 301 SPEPELQLRPYQMEVAQPALEGKNIITICLPTGSGKTRVAVYIAKDHLDKKKKASEPGKVI 360  
Db 301 SPEPELQLRPYQMEVAQPALEGKNIITICLPTGSGKTRVAVYIAKDHLDKKKKASEPGKVI 360

Qy 361 VLVNKVLLVEQLFRKEFQPFLLKWKWVIGLGSQDTQLKISFPEVVKSCDIIISTAQILENS 420  
Db 361 VLVNKVLLVEQLFRKEFQPFLLKWKWVIGLGSQDTQLKISFPEVVKSCDIIISTAQILENS 420

Qy 421 LNNLENGEDAGVQLSDLSLIIIDECHHTNKEAVYNNIMRHYLMOKLKNRNLKKENKPVIP 480  
Db 421 LNNLENGEDAGVQLSDLSLIIIDECHHTNKEAVYNNIMRHYLMOKLKNRNLKKENKPVIP 480

Qy 481 LPQILGLTASPGVGGATKQAKAEEHILKLCANLDAFTIKTVKENLDQLKNOIQLEPCKKFA 540  
Db 481 LPQILGLTASPGVGGATKQAKAEEHILKLCANLDAFTIKTVKENLDQLKNOIQLEPCKKFA 540

Qy 541 IADATREDPFEKLLIMTRIQTYCQMSPMSDFGTQPEQWAIQMEKKAAKKGKGNRKERV 600  
Db 541 IADATREDPFEKLLIMTRIQTYCQMSPMSDFGTQPEQWAIQMEKKAAKKGKGNRKERV 600

Qy 601 AEHLRKYNEALQINDTIRMIDAYTHLETFYNEEKDKKFAVIEDDSDEGGDDEYCDGDEDE 660  
Db 601 AEHLRKYNEALQINDTIRMIDAYTHLETFYNEEKDKKFAVIEDDSDEGGDDEYCDGDEDE 660

Qy 661 DDLKKPKPLKDETDRFLMLTFENNPKMLKRLAENPEYENEKLTKLRLNTIMEQYTRTEESAR 720  
Db 661 DDLKKPKPLKDETDRFLMLTFENNPKMLKRLAENPEYENEKLTKLRLNTIMEQYTRTEESAR 720

Qy 721 GIIIFTKTRQSYALSQWITENEKFAEVGVKAHHLTGAGHSSEFKPMTONEQKEVISKFR 780  
Db 721 GIIIFTKTRQSYALSQWITENEKFAEVGVKAHHLTGAGHSSEFKPMTONEQKEVISKFR 780

Qy 781 GKNLIIATTVAEGLDIKECNIVIRYGLVTNEIAMVQARGRAADESTYVLVAHSGSGV 840  
Db 781 GKNLIIATTVAEGLDIKECNIVIRYGLVTNEIAMVQARGRAADESTYVLVAHSGSGV 840

Qy 841 JEHETYNDFREKMMYKAIIHCYQNMKPEEYAHKILELQMOQSIMEKKMKTKNIAKHYKNNP 900  
Db 841 IERETVNDFREKMMYKAIIHCYQNMKPEEYAHKILELQMOQSIMEKKMKTKNIAKHYKNNP 900

Qy 901 SLITFLCKNCSVLAACSGEDIHVTEKMHVNMTPEFKELYIVRENKALQKKCADYQINGEI 960  
Db 901 SLITFLCKNCSVLAACSGEDIHVTEKMHVNMTPEFKELYIVRENKALQKKCADYQINGEI 960

Qy 961 ICKCGQAWGTMVHKGLDLPCLKIRNFVVFKNINSTKQYKKWVELPITFPNLDSYSECC 1020  
Db 961 ICKCGQAWGTMVHKGLDLPCLKIRNFVVFKNINSTKQYKKWVELPITFPNLDSYSECC 1020

Qy 541 IADATREDPFEKLLIMTRIQTYCQMSPMSDFGTQPEQWAIQMEKKAAKKGKGNRKERV 600  
Db 541 IADATREDPFEKLLIMTRIQTYCQMSPMSDFGTQPEQWAIQMEKKAAKKGKGNRKESVC 600

Query Match 46.2%; Score 2456; DB 15; Length 468;  
 Best Local Similarity 99.1%; Pred. No. 9e-171;  
 Matches 464; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 601 AEHLRKYNEALQINNDTIRMIDAYTHLETFYNEEKDKKFAVIEDDSDEGGDDEYCDGDEDE 660  
 Db 601 AEHLRKYNEALQINNDTIRMIDAYTHLETFYNEEKDKKFAVIEDDSDEGGDDEYCDGDEDE 660

Qy 661 DDLKKPLKLDDETDRFLMTLFFENKMLKRLAENPEYENEKLTKLRLNTIMEQYTRTEESAR 720  
 Db 661 DDLKKPLKLDDETDRFLMTLFFENKMLKRLAENPEYENEKLTKLRLNTIMEQYTRTEESAR 720

Qy 721 GIIIFTKTRQSAVLSQWITENEKFAEVGKAHLLIGAHSSEFKPMTONEQKEVISKPRT 780  
 Db 721 GIIIFTKTRQSAVLSQWITENEKFAEVGKAHLLIGAHSSEFKPMTONEQKEVISKPRT 780

Qy 781 GKIINLLIATTVAEGLDIKECNIVIRGLVTNEIAMVQARGRAADESTYVLVAHSGSGV 840  
 Db 781 GKIINLLIATTVAEGLDIKECNIVIRGLVTNEIAMVQARGRAADESTYVLVAHSGSGV 840

Qy 841 IEEHETVNDFREKOMYKAIHCVQNMKPEEYAHKILELQMSIMEKKMKTKRNIAKHYKNNP 900  
 Db 841 IERETVNDFREKOMYKAIHCVQNMKPEEYAHKILELQMSIMEKKMKTKRNIAKHYKNNP 900

Qy 901 SLITFLCKNCSVLACSGEDIHVIEKMEHVNMTPEFKELLYIVRENKAQDKADYQINGEI 960  
 Db 901 SLITFLCKNCSVLACSGEDIHVIEKMEHVNMTPEFKELLYIVRENKAQDKADYQINGEI 960

Qy 961 ICKCGQAWGTMVVHKGLDLPCLKIRNFVYVFKNNSTKKQYKKWELPITFPNLDYSECCCL 1020  
 Db 961 ICKCGQAWGTMVVHKGLDLPCLKIRNFVYVFKNNSTKKQYKKWELPITFPNLDYSECCCL 1020

Qy 1021 FSDED 1025  
 Db 1021 FSDED 1025

RESULT 6  
 US-10-094-749-2467  
 Sequence 2467, Application US/10094749  
 ; Publication No. US20030219741A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ISOGAI, TAKAO  
 ; APPLICANT: SUGIYAMA, TOMOYASU  
 ; APPLICANT: OTSUKI, TETSUJI  
 ; APPLICANT: WAKAMATSU, AI  
 ; APPLICANT: SATO, HIROYUKI  
 ; APPLICANT: ISHII, SHIZUKO  
 ; APPLICANT: YAMAMOTO, JUN-ICHI  
 ; APPLICANT: ISONO, YUUKO  
 ; APPLICANT: HIO, YURI  
 ; APPLICANT: OTSUKA, KAORU  
 ; APPLICANT: NAGAI, KEIICHI  
 ; APPLICANT: IRIE, RYOTARO  
 ; APPLICANT: TAMECHIKA, ICHIRO  
 ; APPLICANT: SEKI, NAOKI  
 ; APPLICANT: YOSHIKAWA, TSUTOMU  
 ; APPLICANT: OTSUKA, MOTOYUKI  
 ; APPLICANT: NAGAHARI, KENJI  
 ; APPLICANT: MASUHO, YASUHIKO  
 ; TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA  
 ; FILE REFERENCE: 0B4335/0160  
 ; CURRENT APPLICATION NUMBER: US/10/094,749  
 ; PRIOR APPLICATION NUMBER: 60/350,435  
 ; PRIOR FILING DATE: 2002-01-24  
 ; CURRENT FILING DATE: 2002-03-12  
 ; NUMBER OF SEQ ID NOS: 3381  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO: 2467  
 ; LENGTH: 468  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-10-094-749-2467

Query Match 46.2%; Score 2456; DB 15; Length 468;  
 Best Local Similarity 99.1%; Pred. No. 9e-171;  
 Matches 464; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 618 RMIDAYTHLETFYNEEKDKKFAVIEDDSDEGGDDEYCDGDEDE 677  
 Db 618 RMIDAYTHLETFYNEEKDKKFAVIEDDSDEGGDDEYCDGDEDE 677

Qy 678 TLFEENNKKMLKRLAENPEYENEKLTKLRLNTIMEQYTRTEESARGLIFTKTRQSAVLSQW 737  
 Db 121 TLFEENNKKMLKRLAENPEYENEKLTKLRLNTIMEQYTRTEESARGLIFTKTRQSAVLSQW 180

Qy 738 ITENEKFAEVGKAHHLIGAHSSEFKPMTQNEKEVISKFRTGKINLLIATTVAEGLD 797  
 Db 181 ITENEKFAEVGKAHHLIGAHSSEFKPMTQNEKEVISKFRTGRIINLLIATTVAEGLD 240

Qy 798 IECBNIVIRGLVTNEIAMVQARGRAADESTYVLVAHSGSGVIEHETVNDFREKMMYKA 857  
 Db 241 IECBNIVIRGLVTNEIAMVQARGRAADESTYVLVAHSGSGVIEHETVNDFREKMMYKA 300

Qy 858 IHCVQNMKPEEYAHKILELQMSIMEKKMKTKRNIAKHYKNNPSSLITFLCKNCSVLACSG 917  
 Db 301 IHCVQNMKPEEYAHKILELQMSIMEKKMKTKRNIAKHYKNNPSSLITFLCKNCSVLACSG 360

Qy 918 EDIHTVIEKMHVNNTPEFKELLYIVRENKAQDKADYQINGEIICKCGQAWGTMVVHKGL 977  
 Db 361 EDIHTVIEKMHVNNTPEFKELLYIVRENKAQDKADYQINGEIICKCGQAWGTMVVHKGL 420

Qy 978 DLPCLKIRNFVYVFKNNSTKKQYKKWELPITFPNLDYSECCCLFSDED 1025  
 Db 421 DLPCLKIRNFVYVFKNNSTKKQYKKWELPITFPNLDYSECCCLFSDED 468

RESULT 7  
 US-10-408-765A-2031  
 ; Sequence 2031, Application US/10408765A  
 ; Publication No. US20040101874A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ghosh, Soumitra S.  
 ; APPLICANT: Fahy, Eoin D.  
 ; APPLICANT: Zhang, Bing  
 ; APPLICANT: Gibson, Bradford W.  
 ; APPLICANT: Taylor, Steven W.  
 ; APPLICANT: Glenn, Gary M.  
 ; APPLICANT: Warnock, Dale E.  
 ; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION  
 ; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME  
 ; CURRENT APPLICATION NUMBER: US/10/408,765A  
 ; CURRENT FILING DATE: 2003-04-04  
 ; NUMBER OF SEQ ID NOS: 3077  
 ; SOFTWARE: FastSEQ for Windows Version 4.0  
 ; SEQ ID NO: 2031  
 ; LENGTH: 925  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-10-408-765A-2031

Query Match 20.5%; Score 1087.5; DB 16; Length 925;  
 Best Local Similarity 31.5%; Pred. No. 2.5e-70;  
 Matches 328; Conservative 166; Mismatches 361; Indels 185; Gaps 37;

Qy 12 RYLISCFRARVMMIYQVEFVLDYL-T-FLPAEVKEQIORTVATSGMVAEVLLSTLKG 69  
 Db 6 RRSIQAQDQDYRKTLDPYIILSYMAPWFREEV-QYIQAENKNGPMEAATLFLKFLLE- 63

Qy 70 VWHLGWTREFVEALRRTGSPLAARYMNPPELTDLSPSPSENAH--DEYLQLNNLQPTLVD 127  
 Db 64 LQEEGWFRGFLDALDHAG--YSGLYEAEWSW-----FKKIEKLEEYRILLKRLQPEFKT 116

QY 128 KLLVVDLDKCMEEELLTTIEDNRRIAAEENNNGESGVRELLKRVQ--KENWFAFLNVL 185  
 Db 117 RIPTDIISD-LSECCLINECEEILQICSTKGMMAGAEKLVCEULLRSDKENWPKTL--- 171  
 QY 186 RQTGNNEVLQELTGSDCSESNAEIEENLSQVDPQVEEQLLSTTVQPNLEKEVWGMEN--- 242  
 Db 172 -----KLALEKERNKFSELWIVERGKTDQ-----IFYQEDPEC 219  
 QY 243 -NSSESSFADSSVVSSEDTSLAEGSVSCLDESLGHNSNMGSDSGTMGSDEENVAARAS 301  
 Db 220 QNLSENSCPPSEV---SDTNL-----YS 239  
 QY 302 PEPELQLRPPYQMEVAQPALEGKNIICLPTGSKTRVAVYIAKDHLDKRRKASEPGRKVIV 361  
 Db 240 P---FKPRNYQOLELLALPAMKGKNTIICAPRGCGKTFVSLICEHHL-KKFQPQGKGVVF 295  
 QY 362 LVNKVLLVEQ---LPRKEFQPFLLKKWYRVTIGLSDTQLKTSFPEVVKSCDIIISTAQILE 418  
 Db 296 FANQIPVYEQQKSVPESKYFE--RHGYRVTGIGSATAENVVPVEQIVENNDIILITPQILV 352  
 QY 419 NSLNLNLENGEDAGVQLSDFLIIIDECHETTNKEAVNNIMRHYNMQKLKNNRKKENKPV 478  
 Db 353 N---NLKKGTTIP--SLSLIFTLMIFDECHNTSKQHPYNNMFLDKXLGGSS--- 399  
 Qy 479 IPLLPOQILGLTASPGVGGATKQAKAAEHTLKLCANLDAFTIKTVKENLDQLKNQIQLPCKKK 538  
 Db 400 GPLPQVIGLTTASVGWDAKNTDEALDYICKLCASLDASVIATVKHNLEELQQVYKPKQKF 459  
 Qy 539 FAJADATREDPFKE-----KULEIMTRIOTYCCQMSPMSPMSDFGTQPYEQW 581  
 Db 460 FRKVESRISDKFYIILQALMRTDESLAKRICKDLENLSQIQN----REFGTQKYEQW 512  
 QY 582 AJQMEKKA---AKKGNRKERV-----AEHLRKYNEALQINDTIRMIDAYTHLETFYN 631  
 Db 513 IVTVQKACMVFQMPDKDEESRICKALFLYTSHLRKYNDALIISEHARMKDADYLKDFFS 572  
 QY 632 EEKDKKFAVIEDDSDEGGDDEYCDGDEDDDLKKPLKLDDETRFLMLFFENNMKLKRLA 691  
 Db 573 NVRAAGF-----DETEQDL-----TQR-----FEEKLQELESVS 601  
 QY 692 ENPEYENKLTKLRTNTIMEQYTRTEESARGLIIFTKTRQSAYALSQWITENEK--FAEVGV 749  
 Db 602 RDPSNENPKLEDLCFILQEEYHLPETIT-ILFVKTRALVDALKWIEGNPKLSSLKPGI 660  
 QY 750 KAHHLIGAGHSSEFKPMQTQEKEVISKFR--TGGKINLILLIATTVAEGLDIKECNIVRYG 808  
 Db 661 -----LTGRGKTNQNTGMLPAQKCILDAFKASGDFHNTLIATSADEGTDIAOCNLVLYE 716  
 QY 809 LVTNEIAMVQAGRRAADESTYVLAHSGSGVIEHETNDFREKMMYKALHCVNQNMKPEE 868  
 Db 717 YVGNVIKMIQTRGRGRGSKCFLLT-SNAGVIEKEQINMYKEKMMNDSILRQQTWDEAV 775  
 QY 869 YAHKILELQMQSIMEKKMKTKRNIAKHY--KNNPLSLITFLCKNCSVLACSGEDIHVTEKM 926  
 Db 776 FREKTLHIOHT---EKFIRDSQEKPVPDKENKKL---LCRKCKALACYTADVRVIEEC 829  
 QY 927 HHVNMTPEPKELYIVRENKAЛОKKCADYQINGEIIIC--KCGQAWGTMVYHKGDLPLCLK 983  
 Db 830 HYTVLGDAFKECFVSRPHPK--PKQFSSFEKRAKIFCARQNCSSHWDWGIHVVKTFEIPVIK 888  
 QY 984 IRNFVVVFKNNSTKQYKKW 1003  
 Db 889 IESFVVVEDIATGVQTLYSKWN 908  
 RESULT 8  
 US-10-055-475-12  
 ; Sequence 12, Application US/10055475  
 ; Publication No. US20030022855A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Fisher, Paul B.  
 ; APPLICANT: Kang, Dong-Chul  
 ; LENGTH: 487  
 RESULT 9  
 US-10-108-260A-4515  
 ; Sequence 4515, Application US/10108260A  
 ; Publication No. US20040005560A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: HELIX RESEARCH INSTITUTE  
 ; TITLE OF INVENTION: No. US20040005560A1 full length cDNA  
 ; FILE REFERENCE: HI-A0106  
 ; CURRENT APPLICATION NUMBER: US/10/108,260A  
 ; CURRENT FILING DATE: 2002-03-27  
 ; NUMBER OF SEQ ID NOS: 5458  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 4515  
 ; LENGTH: 487

;

TYPE: PRT ; ORGANISM: Homo sapiens ; US-10-108-260A-4515

Query Match 19.6%; Score 1040.5; DB 15; Length 487;

Best Local Similarity 41.1%; Pred. No. 2.6e-67; Matches 234; Conservative 90; Mismatches 161; Indels 85; Gaps 9;

;

QY 306 LQLRPYQMEVAQPALEGKNTIICLPTGSGKTRVAVYIAKDHLDKKKASEPGBKIVLVNK 365

Db 1 MELRSYQWEVIMPALEGKNTIICLPTGAGKTRAAAYVAKRHL---TVDGAKVUVLVR 56

QY 366 VLLVEQLRKEFQPFKKWYRVLIGLSGDTQLKISFPEVYKSCDIIISTAQILENSLL 422

Db 65 LPVYEQQKSVFSKHFE--RLGYKVAGISGATSDTVCVEQIVENSDDIIILTPQILVNCLT 121

QY 57 VHLVTQ-HGEEFRRMLDGRWTVTTLSDGMGPRAGFGHLAR-----95

Db 426 NGEDAGVQLSDFLIIIDECHHTNKEAVYNNIMRHYLMQKLKNNRLLKENKPVIPPLPOIL 485

Db 96 -----CHHTHKDTVNVNIMSQYLELKLQRAQ-----PLPQVL 127

QY 486 GLTASPGVGGATKQAKAEEHILLCANLDAFTIKTVKENLDOLQKQIQEPCKKFAIADAT 545

Db 128 GLTASPGTGGASKLQLDGAINHVLQCLANLDTWCIMSPONCPCQYHQQPCQYNLCHRR 187

QY 546 REDPFKEKLLIMTRIQTYCQMSRMS-DFGTQPYEQWAIQIMEKKAAGKGNRKERVCAEHL 604

Db 188 SQDPFGDLKKLMDQIHDHLEMPELSRKFGTQMYEQQWVVKLSEAALAGLQEQRVYALHL 247

QY 605 RKYNEALQINDTIRMIDAYTHLETFYNEEKKDKKFAVIEDDSDEGGDDEYCDGDEDDEDLK 664

Db 248 RRYNDALLIHDTVRADVDAALQDFYHREHTVKTOIL-----C-----285

QY 665 KPLKLDETDRFLMLFFENNMKMLKRLAENPEYENENKLTKLNRNTIMEQYTRTEESARGITF 724

Db 286 -----AERRLLALFDDRKNELAHLATHGP-ENPKLEMELKILQRQFS-SSNSPRGIFT 336

QY 725 TKTROSAYALSQWITENEKFAEVGVKAHHLIGAGHSSEFKPMTNEQKEVITSKFRGKIN 784

Db 337 TKTROSAAHSSLILWLOQQQGLQTVDIRAQOLLIGAGNSQSTMQDQEVIQKFDGTLN 396

QY 785 LLIATTVAEGLDIKEBCNIVIRYGLVTNEIAMVQARGRAADESTVYLAHSGSGVIEHE 844

Db 397 LLVATSVAAEGLDIPEHNCVNVVRYGLLTNEISMVQARGRAADQSVYAFVATEGSSRELKRE 456

QY 845 TYNDFREKMMYKATHCYQNMKPEEYAHKIL 874

Db 457 LNEALETLMEQAVAAVQKMDQAEYQAKEL 486

;

RESULT 10 US-10-055-475-13

;

Sequence 13, Application US/10055475

;

Publication No. US20030022855A1

;

GENERAL INFORMATION:

;

APPLICANT: Fisher, Paul B.

;

APPLICANT: Kang, Dong-Chul

;

APPLICANT: Gopalkrishnan, Rahul V.

;

TITLE OF INVENTION: USE OF MDA-5 AS AN ANTIVIRAL AND

;

TITLE OF INVENTION: ANTIPROLIFERATIVE AGENT

;

FILE REFERENCE: A34614-A-PCT-USA-A (070050.1921)

;

CURRENT APPLICATION NUMBER: US/10/055,475

;

CURRENT FILING DATE: 2002-01-22

;

PRIOR APPLICATION NUMBER: PCT/US01/06960

;

PRIOR FILING DATE: 2001-02-28

;

PRIOR APPLICATION NUMBER: 09/515,363

;

PRIOR FILING DATE: 2000-02-29

;

NUMBER OF SEQ ID NOS: 17

;

SOFTWARE: FastSEQ for Windows Version 4.0

;

SEQ ID NO: 13

;

LENGTH: 514

;

TYPE: PRT

;

ORGANISM: sus scrofa

;

US-10-055-475-13

;

RESULT 11 US-09-864-761-45372

;

Sequence 45372, Application US/09864761

;

Patent No. US20020048763A1

;

GENERAL INFORMATION:

;

APPLICANT: Penn, Sharron G.

;

APPLICANT: Rank, David R.

;

APPLICANT: Harzel, David K.

;

APPLICANT: Chen, Wensheng

;

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

;

FILE REFERENCE: Aeomica-X-1

;

CURRENT APPLICATION NUMBER: US/09/864,761

;

CURRENT FILING DATE: 2001-05-23

;

PRIOR APPLICATION NUMBER: US 60/180,312

;

PRIOR FILING DATE: 2000-02-04

;

PRIOR APPLICATION NUMBER: US 60/207,456

;

PRIOR FILING DATE: 2000-05-26

;

PRIOR APPLICATION NUMBER: US 09/632,366

;

PRIOR FILING DATE: 2000-08-03

;

PRIOR APPLICATION NUMBER: GB 24263,6

;

PRIOR FILING DATE: 2000-10-04

;

PRIOR APPLICATION NUMBER: US 60/236,359

;

PRIOR FILING DATE: 2000-09-27

;

PRIOR APPLICATION NUMBER: PCT/US01/00666

;

PRIOR FILING DATE: 2001-01-30

;

PRIOR APPLICATION NUMBER: PCT/US01/00667

SOFTWARE: PatentIn Ver. 3.0  
 SEQ ID NO 5088  
 LENGTH: 154  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: MISC\_FEATURE  
 LOCATION: (11)  
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 NAME/KEY: MISC\_FEATURE  
 LOCATION: (17)  
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 NAME/KEY: MISC\_FEATURE  
 LOCATION: (18)  
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 NAME/KEY: MISC\_FEATURE  
 LOCATION: (23)  
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 NAME/KEY: MISC\_FEATURE  
 LOCATION: (25)  
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 NAME/KEY: MISC\_FEATURE  
 LOCATION: (26)  
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 NAME/KEY: MISC\_FEATURE  
 LOCATION: (28)  
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 NAME/KEY: MISC\_FEATURE  
 LOCATION: (75)  
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 US-10-106-698-5088  
 Query Match 14.7%; Score 782; DB 14; Length 154;  
 Best Local Similarity 94.8%; Pred. No. 3.5e-49;  
 Matches 145; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
 Qy 873 ILELOMQSIMEKQMKTKRNLAHKYKNNPSLITFLCKNCSVLACSGEDIHVIEKMHVYMT 932  
 Db 2 ILELOMQSIXEKKMKKXXNLAXHXXNXPSSLITFLCKNCSVLACSGEDIHVIEKMHVYMT 61  
 RESULT 13  
 US-10-264-049-2244  
 CURRENT APPLICATION NUMBER: US/10/264,049  
 ; Sequence 2244, Application US/10264049  
 ; Publication No. US20040005579A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Birse et al.  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
 ; FILE REFERENCE: PA133P1  
 ; CURRENT APPLICATION NUMBER: US/10/264,049  
 ; PRIORITY FILING DATE: 2002-10-04  
 ; PRIOR APPLICATION NUMBER: PCT/US01/18569  
 ; PRIOR FILING DATE: 2001-06-07  
 ; PRIOR APPLICATION NUMBER: US 60/209,467  
 ; PRIORITY FILING DATE: 2000-06-07  
 ; NUMBER OF SEQ ID NOS: 4360  
 ; SOFTWARE: PatentIn Ver. 3.1  
 ; SEQ ID NO 2244  
 ; LENGTH: 154  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: MISC\_FEATURE  
 ; LOCATION: (11)  
 OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids

RESULT 12  
 US-10-106-698-5088  
 ; Sequence 5088, Application US/10106698  
 ; Publication No. US20030109690A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ruben et al.  
 ; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide  
 ; FILE REFERENCE: PA005P1  
 ; CURRENT APPLICATION NUMBER: US/10/106,698  
 ; PRIOR APPLICATION NUMBER: PCT/US00/26524  
 ; PRIORITY FILING DATE: 2000-09-28  
 ; PRIOR APPLICATION NUMBER: US 60/157,137  
 ; PRIOR FILING DATE: 1999-09-29  
 ; PRIOR APPLICATION NUMBER: US 60/163,280  
 ; PRIOR FILING DATE: 1999-11-03  
 ; NUMBER OF SEQ ID NOS: 8564

FEATURE: MISC\_FEATURE  
 LOCATION: (17)  
 OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids

FEATURE: MISC\_FEATURE  
 LOCATION: (18)  
 OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids

FEATURE: MISC\_FEATURE  
 LOCATION: (23)  
 OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids

FEATURE: MISC\_FEATURE  
 LOCATION: (25)  
 OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids

FEATURE: MISC\_FEATURE  
 LOCATION: (26)  
 OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids

FEATURE: MISC\_FEATURE  
 LOCATION: (28)  
 OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids

FEATURE: MISC\_FEATURE  
 LOCATION: (75)  
 OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
 US-10-264-049-2244

Query Match 14.7%; Score 782; DB 15; Length 154;  
 Best Local Similarity 94.8%; Pred. No. 3.5e-49;  
 Matches 145; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 873 ILELOMQSIMEKKMKTKRNIAKHYKNNPSSLITFLCKNCNSVLACSGEDIHVIEMHHVNM 932  
 Db 2 ILELOMQSIXEKKMKTKRNIAKHYKNNPSSLITFLCKNCNSVLACSGEDIHVIEMHHVNM 61

Qy 933 PEFKELYIVRENKAQKCADYQINGEIIICKCQGAWGTMVHKGLDPCLKIRNFVVFK 992  
 Db 62 PEFKELYIVRENKAQKCADYQINGEIIICKCQGAWGTMVHKGLDPCLKIRNFVVFK 121

RESULT 15  
 US-10-264-049-2303  
 ; Sequence 2303, Application US/10264049  
 ; Publication No. US20040005579A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Birse et al.  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
 ; FILE REFERENCE: PA133P1  
 ; CURRENT APPLICATION NUMBER: US/10/264,049  
 ; CURRENT FILING DATE: 2002-10-04  
 ; PRIOR APPLICATION NUMBER: PCT/US01/18569  
 ; PRIOR FILING DATE: 2001-06-07  
 ; PRIOR APPLICATION NUMBER: US 60/209,467  
 ; PRIOR FILING DATE: 2000-06-07  
 ; NUMBER OF SEQ ID NOS: 4360  
 ; SOFTWARE: PatentIn Ver. 3.1  
 ; SEQ ID NO 2303  
 ; LENGTH: 447  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-10-264-049-2303

Query Match 10.6%; Score 563.5; DB 15; Length 447;  
 Best Local Similarity 33.0%; Pred. No. 1.6e-32;  
 Matches 155; Conservative 75; Mismatches 170; Indels 69; Gaps 16;

Qy 553 KLEIMTRIQTYCQMSSPMSDFGTQPYQWAIQMEKKA---AKGNRKRVC---AE 602  
 Db 13 KDLNLSQLQIN-----REFGTQKYEQWIVTQKACMFQMPDKDEESRICKALFLYTS 65

Qy 603 HLRKXNEALQINDTIRMIDAYTHLETFYNEEKDKKFAVIEDSDSDEGGDDEYCDGDEDEDD 662  
 Db 66 HLRKXNEALQINDTIRMIDAYTHLETFYNEEKDKKFAVIEDSDSDEGGDDEYCDGDEDEDD 662

Query Match 11.6%; Score 617; DB 14; Length 416;  
 US-10-055-475-14  
 ; LENGTH: 416  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-10-055-475-14

Qy 663 LKKPLKLDETDRLMTLFFENNKKMLKRLAENPEYENEKLTKLRRNTIMEQYTRTEESARGI 722  
 Db 108 L-----TQR-----PEEKQLELSVSRDPSNNENPKLEDLCFTLQEEYHLPETIT-I 153  
 Qy 723 IFTKTROSAYALSQMITENEK--FAEVGVKAHHLIGAGHSSEFKEMTONEQEVISKFR- 779  
 Db 154 LFVKTTRALVDALKNWIEGNPKLSFLKPGI----LTGRGKTNQNTGMLPAQKCILDAFKA 209  
 Qy 780 TGKINLLIATTVAEEGLIKECNIVIRYGLVTNEIAMVQARGRAADESTYVLAHSGSG 839  
 Db 210 SGDHNILIASTVADEGIDIAQCNLVILYYGVNVIKMIQTRGRGARGSKCFLLT-SNAG 268  
 Qy 840 VIEHETVNDFREKMMYKAIHCVQNMKPEEYAHKILELQMQSIMEKMKTKRNIAKHY--K 897  
 Db 269 VIEKEQINMYKEKMMNDSLRLQTWDEAVFREKILHITH---EKFIRDSQEKPKPVPDK 325  
 Qy 898 MNPSLITFLCKNCVULACSGEDIHVIEKMHVDMTPEFKELYIVRENKALQKCADYQIN 957  
 Db 326 ENKCL---LCRKCKALACYTADVRVTEBCHYTVLGDAFKECFVSRPHPK-PKQFSSFEKR 381  
 Qy 958 GEIIC---KCGQAWGTMVHKGLDLPCLKIRNFVVFKNSTKKQYKKW 1003  
 Db 382 AKIFCARQNCSDWGIHVKYKTFEIPVTKIESFVVEDIATGVQTLYSKW 430

Search completed: June 2, 2004, 19:18:02  
 Job time : 57 secs

GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model  
 Run on: June 2, 2004, 19:07:33 ; Search time 27 Seconds  
 (without alignments)  
 3651.715 Million cell updates/sec

Title: US-09-515-363C-2  
 Perfect score: 5311  
 Sequence: 1 MSNGYSTDENFRYLISCFRA.....LPITFPNLDYSECCCLFSDED 1025  
 Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% ; Maximum Match 100% ; Listing First 45 summaries

Database : PIR\_78:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	687	12.9	1037	2	T32534		hypothetical prote
2	654	12.3	811	2	T30968		helicase homolog C
3	642.5	12.1	956	2	A88708		protein C01B10.1 [
4	516	9.7	398	2	T46312		hypothetical prote
5	499	9.4	1134	2	T20332		hypothetical prote
6	453	8.5	752	2	D75219		ATP-dependent RNA
7	441.5	8.3	650	2	D71203		probable ATP-dependent RNA
8	417	7.9	741	1	A69432		ATP-dependent RNA
9	402.5	7.6	1374	2	S62524		probable RNA helic
10	392.5	7.4	784	2	H64487		elF-4A family prob
11	350	6.6	1822	2	S44849		X12H4.8 protein -
12	342	6.4	738	1	E69055		ATP-dependent RNA
13	313.5	5.9	784	2	D84386		ATP-dependent RNA
14	303.5	5.7	993	2	S48436		probable RNA helic
15	284	5.3	502	2	T31323		ATP-dependent RNA
16	278.5	5.2	839	2	T39190		probable ATP-dependent RNA
17	223	4.2	557	2	G69958		SNF2 helicase homo
18	218	4.1	1979	2	C71622		hypothetical prote
19	217.5	4.1	2663	1	S28261		centromere protein
20	214.5	4.0	1790	2	S67593		transport protein
21	209.5	3.9	449	2	C71156		probable helicase
22	205	3.9	453	2	B75006		DNA repair protein
23	205	3.9	967	2	D72308		conserved hypothet
24	204.5	3.9	2163	2	S50675		pre-mRNA splicing
25	204	3.8	1939	2	T18372		repeat organellar
26	201.5	3.8	1063	2	T38420		probable DNA helic
27	198.5	3.7	2253	2	T30336		nuclear/mitotic ap
28	197.5	3.7	1163	2	G97236		ATPase involved in
29	195	3.7	1531	2	T48946		hypothetical prote

30	195	3.7	3259	1	A56539	giantin - human
31	194.5	3.7	986	2	E90596	restriction-modifi
32	194.5	3.7	1256	2	T26101	hypothetical prote
33	194	3.7	1039	2	E72734	hypothetical prote
34	193.5	3.6	663	2	G89848	exinuclelease ABC su
35	193.5	3.6	2017	1	A36014	myosin heavy chain
36	193.5	3.6	2057	2	S61477	giantin - human
37	193.5	3.6	3225	2	I52300	DNA helicase relat
38	193	3.6	715	2	B75135	restriction modifi
39	193	3.6	986	2	H90565	CG1 protein - huma
40	192.5	3.6	718	2	E71100	HsDR1 protein - My
41	192.5	3.6	1300	2	I53799	kinectin 1 - human
42	192	3.6	986	2	S49394	ATP-dependent RNA
43	192	3.6	1356	2	S32763	hypothetical prote
44	191.5	3.6	943	2	F69543	
45	191	3.6	1922	2	T00637	

#### ALIGNMENTS

RESULT 1						
T32534						
hypothetical protein F15B10.2 - Caenorhabditis elegans						
C;Species: Caenorhabditis elegans						
C;Date: 29-Oct-1999 #sequence_change 29-Oct-1999						
C;Accession: T32534						
R;Wohldmann, P.; Murray, J.						
submitted to the EMBL Data Library, December 1997						
A;Description: The sequence of C. elegans cosmid F15B10.						
A;Reference number: Z21186						
A;Accession: T32534						
A;Status: preliminary; translated from GB/EMBL/DDBJ						
A;Molecule type: DNA						
A;Residues: 1-1037 <WOH>						
A;Cross-references: EMBL:AF036696; PIDN:AB88350.1; GSPDB:GN00022; CESP:F15B10.2						
A;Experimental source: strain Bristol N2; clone F15B10						
C;Genetics:						
A;Gene: CESP:F15B10.2						
A;Map position: 4						
A;Introns: 30/3; 89/1; 111/2; 218/1; 258/2; 288/1; 343/3; 370/3; 400/2; 443/1; 520/1; 558						
Query Match Score 12.9% ; DB 2; Length 1037;						
Best Local Similarity 25.0% ; Pred. No. 1.4e-27;						
Matches 251; Conservative 169; Mismatches 408; Indels 174; Gaps 35;						
Qy	101	DLPSPSFENAHD	--YLQLQNLQOPTLVDRKLIVYDVKCMEE	--	142	
Db	63	DLKNSILSNADDERLYKIDMTYQL-TYLPKCTVHKLLN-CSNREVKLSDFHYILDHFEGF	120			
Qy	143	LITIEDRNRIAAAENNGNEGVRELLKRIVQK	--ENWFSAFLN--VLRQTG--NNNELVQ	195		
Db	121	LRPIEPKVVLAYLDSDYPQYIDAVAVLKEIERNEEDNQDSDFI	KLILRTVPLLGQEQAVY	180		
Qy	1196	EL--TGSDCSESNNAE1ENLSQLDGPQVEEQLLSTTVQPNLEKEYWGMEN	--N	243		
Db	1181	DIMYTISEKSSNNLDVE--AKQFIA--KVLRLKNDGFLRFLYQIIN	221			
Qy	244	SSESSFADESSVVSSESDTSILAEGSVSCLDLSIGHNS	--NMGSDSGTMGSDSDE--ENV	296		
Db	2222	ASRQLNQRLIYICPVHESATEMMVYLTGAAALNTNRYMINRVDNTVQENSTPRVIESV	281			
Qy	297	AARRASPEPELQLRPYOMEVAQPALEGKVNLLICLPLTGSKTRAVVYI	KDQKASEP	356		
Db	282	RQRTHRQRLCLRNQEEQELCQVALQGKNTIVTAPTSQGKTVIAANL	IKHFEFSRSSLIVATPQ	341		
Qy	357	GKVIVLV-NKVLVLEQLPFRKEFQPFKKWRYVIGLSGDTQKLKISEPEVVKSCDII	LISTAQ	415		
Db	342	FKALFMTPNMILNQQ--AASISYYLDHVHTQIIQGSD--NVPTRNVIQSOKDLIVATPQ	397			
Qy	416	ILENSLLNLENGEDAVQ	--LSDFLSLLIIDECHHTNKEAVYNNMRHMLQKLNRR	470		
Db	398	MIVNLNCNEHRNSLDDESRLDQFFLSTFTIIFDECHNTVKNSPYSNTMREY--HYLKXNMG	455			



Db	21	LESIYRDPKEKCFSELLPLGKIDELKTHS-ENSQEFSKQYQDLKKNSNLIAADP----E	75	C; Accession: T46312 R; Duesterhoeft, A.; Lauber, J.; Mewes, H.W.; Weil, B.; Wiemann, S. submitted to the Protein Sequence Database, January 2000
QY	129	LLYRDVL-----DKCMEEELLTTIEDRNRRIAAEENNNGESVREILKRIVKENWFAF	181	A; Reference number: Z23036
Db	76	RLYKDIMAYLQSNEFPKCILHDVLNYSGRN-VKLSEFRYILDHLEGFLRLIEPK----VV	129	A; Accession: T46312 A; Status: preliminary A; Molecule type: mRNA A; Residues: 1-398 <AAA> A; Cross-references: EMBL:AU137608 A; Experimental source: adult testis; clone DKFZp434J1111 C; Genetics: A; Note: DKFZp434J1111.1
QY	182	LNVLROTGN-NELVQELTGSDCSESNAELENLSQVDGPQVEEQLLSTVQPNLEKEVWGM	240	
Db	130	LNVLDSYPOYSEAVKLL--KDEIDRHEEDNN----DPEFIKKLILRTVPLLGQAVYDI	182	
QY	241	ENNSSESSFADESSVVSSESDTSLA-EGSVSCLDESGLHNSNMGSDSDEENVAAR	299	
Db	183	MYTIWEKAQMSTSIIINIRVDSTIAPENSASRL-----VIESVRQR	221	Query Match 9.7%; Score 516; DB 2; Length 398; Best Local Similarity 33.3%; Pred. No. 2.e-19; Matches 141; Conservative 68; Mismatches 157; Indels 58; Gaps 14;
QY	300	ASPEPELQLRPyQMEVAQPALEGKNIICLPTGSKTRVAVYIAKDHLDKKKKASEPGKV	359	
Db	222	IHIQROFCLRNYQEEICQVALQGKNTIVTAPPGSKTVAANILKEHFESR--SSEG-	276	QY 594 NRKERVCA----AEHLRKYNEALQINDTIRMIDAYTHLETFYNEEKDKKFAVIEDDSDE 647 2 DEESRICKALFLYFTSHLRKYNDAIISEHARMKDALDLKDFFSNVRAAGF----- 52
QY	360	IVLVNKVLLVEQLPFRKEFQPFPLKWRVIGLSGDTQLKISFPEVVKSCDIIISTAQILEN	419	
Db	277	-KRFKTQTIQGSD--NVPTRNVIQSKDLIVATPQMIWN 311		QY 648 GGDEYCDGDEDDEDDLKKPKLKDDETDRFLMTLFFENNIKMLKRLAENPEYENELTKLRLNT 707 53 -----DETEQDL----TQR-----TQR-----PEEKLQEELESVSRDPSENPKLEDCFCI 90
QY	420	SLLNLENGEDAGVQ----LSDFSLIITIDECHHTNKEAVYNNNIMRHYLMQKLKNRILK	472	Db 91 LQBEYHLLNPETIT-ILFVKTREALVDAALKNWIEGNPKLSPFLKPGI----LTGRGKTNQNTG 145
Db	312	-LCN-EHRDPDDEYPPEQFFLSTETLIFPDECHNTVKNSPYSNVMREY--HYLKNMGMN	367	QY 708 IMEQYTRTEESARGLIIFTKTRQSAYALSQWITENEK--FAEVGVKAHHLIGAGHSSEFPKP 765
QY	473	KENKPVIPLPQIQLGTTASPGVGGATKQAKAEEHILKLCANLDAFTTITKVKENLDQLKRNQI	532	Db 914 MTLAQKCLLDAFKASGDNHNLIASTSVADEGIDIAQCNLVLVILYEEYGVNVNIQTRGRGR 205
Db	368	PEGH---SFPQIIGLTASLGTGDKKKNCMQVRSYTAGLCANMDVKEILSIVKDNLLELLDHN	424	QY 766 MTQNEQKEVISKFR-TGKLNLLIATTVAEGGLDIKECNIVIRYGLYTNEIAMQARGRAR 824
QY	533	QEPCKKFAIADATREIDP----FKEKLLIEIMTRIQTQYQMSDFGTQP-----Y	578	Db 146 ARGSKCFLLT-SNAGVIEKEQIANYKEKMMANDSILRLQTWDEAVPREKILHIQTH--EK 261
Db	425	FVVTDDQVSFCERSNDGPIEMFTKRLKQMMOEVEDLIRTLKNEPTVKEIIPPTDKEHNRY	484	QY 825 ADESTYYLVAHSGSGVIEETVNDFREKMMYKAIIHCVQNMKPEEYAHKILELQMOSIMEK 884
QY	579	EQWAIQMEKKAQKGNBKERVCAEHLRKYNEALQINDTIRMIDAYTHLETFYNEEKDKKF	638	Db 206 ARGSKCFLLT-SNAGVIEKEQIANYKEKMMANDSILRLQTWDEAVPREKILHIQTH--EK 261
Db	485	ENWISNQRNCVSLAGSRNKTLLIEVL-----DVLKVQYMFQSQMKIIFWNYFRKYRF	534	QY 885 KMKTAKHYY--KNNPSLITFLCKNCVLAACSGEDIHVIEKMHMHVNNTPEFKELYIYR 942
QY	639	AV--IEDDSDEGGDDEYCDGDEDDEDDLKKPKLKD----ETDRFLMTLFFENNIKMLK	688	Db 262 FIRDSQEKKPKPVPDFKENKKL---LCRKCKALACYTADVRVIEECHYTIVLGDAFKECFVSR 31.8
Db	535	GFKRIFQFQDCFYALSYNINFNPENPEVALKTYKEKLGPERIRNFTDN--MNLWDN--CHR	590	QY 943 ENKALQKKCADYQINGELIC--KCGQAWGTMWVKGLDLPLCKLIRNFVVFKNNNSTKKQ 999
QY	689	RRAENPEYENKEKLTKLKRNTIMEQYTRTEESARGLIIFTKTRQSAYALSQWITENEKFAEVG	748	Db 319 PHPK-PKQFSSFEKRAKIFCARQNCNSHDWGIHVKYKTFEIPVIKIESFVVVEDIATGVQTL 377
Db	591	ELVGIGSAENPMIARTVQFILDQNEQTSD-FRAIIIVRTKEADEFLNYVL--NDRLHELG	647	QY 1000 YKKW 1003
QY	749	VKAHHLIG--AGHSSEFKPMTQNEQREVISKFRTGKINLLIATTVAEGGLDIKECNIVI	805	Db 378 YSKW 381
Db	648	IKSDWMSGQKKSTASSADISASKQKOMEKLFMDAGENQILVSTSYVAEGLDIPECSLVI	707	RESULT 5 T20332
QY	806	RYGLVTNEIAMVQARGRADAESTYVLAHSGGVIEH--ETVNDDFREKMMYKAIIHCVCQV	863	hypothetical protein D2005.5 - Caenorhabditis elegans
Db	708	KYNYATNETAHVQRGRBARNSKCVLITNS---IAHVQESNNLAKENLMTETISLION	764	C; Species: Caenorhabditis elegans C; Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
QY	864	MKPE----EYAHKLLQELQMOSIMEKKMKTKRNIAKHYKNNPSLITFLCKNCVLAACSG	917	C; Accession: T20332 R; Wilkinson, J. A; Reference number: Z19258 A; Accession: T20332 A; Status: preliminary; translated from GB/EMBL/DDBJ A; Molecule type: DNA A; Residues: 1-1134 <WIL>
Db	765	SPGEFRQCVDEESNKKWPRQEDTDKQARIKEQI----NRNIVYKIVCMKCDTVLCTN	819	A; Cross-references: EMBL:Z79752; PIDN: CAB02082.1; GSPDB:GN00019; CESP:D2005.5 A; Experimental source: clone D2005 C; Genetics: A; Gene: CESP:D2005.5 A; Map position: 1 A; Introns: 33/2; 91/2; 116/3; 144/1; 361/1; 421/3; 482/2; 591/1; 684/3; 757/1; 95
QY	918	EDIHVIEKMHMHVNNTPEF----KELYIVRENKAQKKCADYQINGELIC--KCGQAWG	970	RESULT 4 T46312 hypothetical protein DKFZp434J1111.1 - human (fragment)
Db	820	KDLR-SKNTQYIVCNPBGFWSLVRRIPLQEQRASNK---FNSTGGTIECLGERGSXLGQ	874	C; Species: Homo sapiens (man) C; Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
QY	971	MMVHKGLDLPCLKIRNFVVVFKNNSTKKQYKXWEL----PITFPNL 1014		
Db	875	LIDVNTVNLPLCKVKSILLIESNTERILVQWNILDEHFTPPTLXQRD 924		

QY	294	ENVAARASPEP-----ELQLRPYOMEVAQPALEGKRNITICLPTGSKTRVAVY	341	A;Experimental source: strain Orsay
Db	345	EDVCTTIRQPKDSGYNPDAVVTTELRLRTYQEEVQPALEGKNCVIVAPTGSKTRVAVY	404	C;Genetics: A;Gene: PAB0190 C;Superfamily: ATP-dependent RNA helicase eIF-4A
QY	342	IAKDHDLKKKASEPQGKIVLVLNVKLLVEQLFRKEFQPFLLKKWVRRVIGLSCDTQLKTSFP	401	F;33-40/Region: nucleotide-binding motif A (P-loop) #status atypical
Db	405	AALKHIEERTSQGKPSRVVLLVPKIPLYGQO-KDRFLKYGNGMYEVNGFHG-SESSVSGT	462	F;131-136/Region: nucleotide-binding motif B F;135-138/Region: DEAD/H motif
		Query Match 8.5%; Score 453; DB 2; Length 752;		
		Best Local Similarity 25.3%; Pred. No. 9.4e-16;		
		Matches 124; Mismatches 259; Indels 156; Gaps 28;		
		Matches 183; Conservative 306		
QY	402	--EVVKSCDIIISTAQILENSLNLLENEDAGVQLSDFSLITIDECHHTNKEAVYNNIM	458	QY;33-40/Region: nucleotide-binding motif A (P-loop) #status atypical
Db	463	GRRDEVIAHTSVVMPQOLINMLOSVRQNE--RLYVSDFSMMIFDEVHKAAKNHPY--VL	518	F;131-136/Region: nucleotide-binding motif B F;135-138/Region: DEAD/H motif
QY	459	RHYLMQKLNKNRLLKKENKPKVPIPLPQILGLTASP--VGAGATKQAKAEEHLKLCANLDAF	516	Query Match 8.5%; Score 453; DB 2; Length 752;
Db	519	INQMVQEWKYEK-----PQIIGLTASLSVKVDGQKDENQMLNDIYNMLALINAP	567	Best Local Similarity 25.3%; Pred. No. 9.4e-16;
QY	517	TIKTV--KENLDQIQLNQIQEPCKKFAIADATREDPKE-----KLEIM-----	558	Matches 124; Mismatches 259; Indels 156; Gaps 28;
Db	568	HLSTITRQSSIDELNEHVGKPDDSVELCLPAKENILRDYTERYLNHAHGKFLEELASMSK	627	Matches 183; Conservative 306
QY	559	-----TRIQTYCQMSPMSDFGTQPYE-----QWAQMEKKAAKKGNRKERVCAE	602	QY;33-40/Region: nucleotide-binding motif A (P-loop) #status atypical
Db	628	STGRNNTIPENMINTFEKKNQPKN-----YEYDSSLQGTIQLNPKNSQTVAK	681	F;131-136/Region: nucleotide-binding motif B F;135-138/Region: DEAD/H motif
QY	603	HLRKYNEALQINDTIRMIDAYTHLETFYNEEKDKKFAVIEEDDSDEGGDDEYCDGDEDEDD	662	Query Match 8.5%; Score 453; DB 2; Length 752;
Db	682	YMKVYLEARGTIVDLMAMPAMVAFKYE-----KAIGKLNESHSE--TVEYSTETIKDHDHT	731	Best Local Similarity 25.3%; Pred. No. 9.4e-16;
QY	663	LKKPLKLKLDETDREFLMTLFFENNMKMLKRLAENPEYENEXKLTKLRNTNTIMEQYTRTEESARGI	722	Matches 124; Mismatches 259; Indels 156; Gaps 28;
Db	732	LKOTIQSVPE-----IVLRL-----KKYT--HQSPVHQFGNYGEQMVGY	769	Matches 183; Conservative 306
QY	723	IFTKTKTROSAYALSQWITENEKFAEVGVKAHHLIGAGHSSEFKPMTQNQEKEVLSKERTGK	782	QY;33-40/Region: nucleotide-binding motif A (P-loop) #status atypical
Db	770	VIGTNKQGAV-----QTSQQQLTIDKETNGR	797	F;131-136/Region: nucleotide-binding motif B F;135-138/Region: DEAD/H motif
QY	783	INLLIATTVAABEGLDIKECNIVTRYGLYT-NEIAMIQVQARGRADESTYVLVPHSGSGVI	841	Query Match 8.5%; Score 453; DB 2; Length 752;
Db	798	LKVIVATSVVVBEGLDTVTACNLIIKYNCSSGSAIQLVQQGRARAKNSRSVLLSYK-SSIN	856	Best Local Similarity 25.3%; Pred. No. 9.4e-16;
QY	842	EHETVNDFREKOMYKAITHCVQNMKPEEYAHKILELQMOSIMEKKMKTKRNIJAKHYKNPNS	901	Matches 124; Mismatches 259; Indels 156; Gaps 28;
Db	857	ETETNALISSEKYMRLCVKKITENGKEKOLAAEVKRAEILNAERKRNLEBQNLRLRHENK	916	Matches 183; Conservative 306
QY	902	LITFLCKNCVSLACSGEDIHV-----IEKOMHAYNMTPEFKELYIVRENKALQ	948	QY;33-40/Region: nucleotide-binding motif A (P-loop) #status atypical
Db	917	TYKLMCSNCSKEFC--KSIYIKKVFNSNYMFDPSSVWRFLHV----ESVETFI----KCLK	966	F;131-136/Region: nucleotide-binding motif B F;135-138/Region: DEAD/H motif
QY	949	--KKC--ADYQIN-----GEIIC-KCG	965	Query Match 8.5%; Score 453; DB 2; Length 752;
Db	967	ITWKCRRIADYQIAEFPNFAFRQLTERFLCNFQMFQKRVSKYLEDNQPLSLDIKCFHCK	1026	Best Local Similarity 25.3%; Pred. No. 9.4e-16;
		RESULT 6		
QY	966	QAWGTMWVHKGLDPLCILKIRNFVIVVFKANSTTKQYKKWVEL	1006	Db;628 VH 629
Db	1027	LDVGRAYKIRGTYLPQLSVKALTEVQESDYSSTMTKAKWSDV	1067	Db;628 VH 629
		RESULT 7		
D75219		ATP-dependent RNA helicase, eif-4a family PAB0190 - Pyrococcus abyssi (strain Orsay)		
C;Species: Pyrococcus abyssi		Db;628 VH 629		
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 10-May-2001				
C;Accession: D75219				
R;Anonymous Genoscope				
A;Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome structure				
Submitted to the EMBL Data Library, July 1999				
A;Reference number: A75001				
A;Status: Preliminary				
A;Molecule type: DNA				
A;Residues: 1-752 <KAW>				
A;Cross-references: GB:AJ248283; GB:AL096836; PID: CAB49203.1; PID: g545771				

DNA Res. 5, 55-76, 1998  
 A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic archaeon  
 A;Reference number: A71000; MUID:98344137; PMID:9679194  
 A;Accession: D71203  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-650 <RAW>  
 A;Cross-references: GB:AP000007; NID:932236134; PID:BAA31019.1; PMID:93258336  
 A;Experimental source: strain OT3  
 A;Note: this accession replaces an interim accession for a sequence replaced by GenBank  
 C;Genetics:  
 A;Gene: PH1896  
 C;Keywords: ATP; nucleotide binding; P-loop  
 F;33-40/Region: nucleotide-binding motif A (P-loop)  
 F;131-136/Region: nucleotide-binding motif B  
 F;135-138/Region: DEAD/H motif

Query Match 8.3%; Score 441.5; DB 2; Length 650;  
 Best Local Similarity 25.9%; Pred. No. 3e-15;  
 Matches 185; Conservative 122; Mismatches 261; Indels 147; Gaps 29;

QY 306 LQLRPYQOMEVAQPALEGKNNIICLPTGSKTRVAVYIAKDHLDKKKAKASEPGKVIVLV-N 364  
 Db 11 LQPRLYQ-ELIYAKCKERNCLIVLPTGLGKTIIAMMIADYRLDK---YGGKVLMLAPT 64  
 QY 365 KVLLVE--QLERKEFQQFLKKWYRVLGIGLSDTQLKISFPEVVKSK--CDIIISTAQILEN 419  
 Db 65 KPLVLQHAETFRKFFTLPPPEK---IVALTGE---ISPNEVRRAWAKVVIATPQTIEN 117  
 QY 420 SLLNLLENGEDAGVQLSDFSLLIIIDECHHTNKEAVYNNIMRHYLMQKLKNNRLLKRENKPVI 479  
 Db 118 DLL---VGRISLEDTLIIIFDEAHRAVGNYAYVYIAKEYLHQ-AKN---159  
 QY 480 PLPQILQLTASPVGGGATQOAKAEEHILKLCANLDAFTIKTKVKENLDQLENQIQEPCKKF 539  
 Db 160 -PHVIGLTASPG-----STPEKIMEVLRNLGIEHIEYRSSENSPDVKPYVQG--IKF 207  
 QY 540 AIAADATREDPFKE--KLLLEIMTRIQLTYCQMSDFTQTOPYEQWAIQMEKKAAKKGNR-- 595  
 Db 208 EWWKVELPELYKEVRKLLRDMLR---DSLKPLAEGL--LDSTSADIPRKETILRAGQII 261  
 QY 596 -KERVCAEH-LRK---YNEALQINDTIRMDAY----THLETFYNEEEKDKKKFAVIED 643  
 Db 262 NEEEMAKGNHDLRKLJLFFHAMALKHHAIELLETQGLSALRYYKLRLYEEAK---312  
 QY 644 DSDEGGDDEYCDGDEDDEDLKKPLKLDDETDRFLMLPFENNKLKRLAENPEYENKEKLT 703  
 Db 313 -----AGSTRASKELFLDKRMKKALALLIQ-----AKELGTDHPKMEV 350  
 QY 704 LRNTIMEQYTRTRTEESARGILIFTKTKRQSAYALSQWITENEKFAEVGVKAHLLIGAGHSSEF 763  
 Db 351 LKGGLIREQLKR-KENSKLIVFTNYTETAKKV-----EELMDGKIKARRFVGQASREND 403  
 QY 764 KPMTONEQKEVYISKFRGKINLILLATTVAEGGLDIKECNIVIRGLVTNEIAMVQARGR- 822  
 Db 404 RGMSQKEQKLLDILDAFARSEFNVLVATSVGEGLDVPEVDLVIFYEPVPSAIRSVQRRGRT 463  
 QY 823 ARADESTYVLAHSGSGVIEHETVNDFREKMYKAITHCYQNMKPEEYAHKILELQMOISIM 882  
 Db 464 GRQKPGRUVILMAQGTRDEAYWSSRQEKEKIMRETIRMSQM---VRKERQQLSLEYV 518  
 QY 883 EKKMKTIKRNTIAKHYKNNPNSLITFLCKNCNSVLAACSGEDITHTEKMHHVNMTPEFKEL--- 938  
 Db 519 KREVNEE--IKEEKKEETGGIKVYIDSRELR-----EVVKKLKLTGKIEVRTLGVGD 569  
 QY 939 YIVRENKAQKKCADYQING-----ELICKCGQAWGTMTH 974  
 Db 570 YIVSEDVAIERSANDFIQSIIIDGRLFDQVKRLKEAYPRPVILLEGGQLYGIRNTVH 624  
 QY 914 ACSGEDIHVIEKMHVNMTPFKEFEL---YIVRENKAQKKCADYQINGEI 960

RESULT 9

Db 532 S-----EVVKHLREIGAKIEIRNLLEVADYVSDRVAVERKTFEDFLNSII 576

Db 297 KALGKQLSD-----DEE-----LAIDRLKIFVED--WKNNNKYSNDNGPRIPVFD 337

QY 698 N----EKLTKLIRNNTIMEQYTRTEESARGIIIFTKTROSAYALSQWITENEKFAEVGKAAH 753

C:Species: Schizosaccharomyces pombe

C:Date: 16-May-1996 #sequence revision 13-Mar-1997 #text\_change 17-Nov-2000

C:Accession: T39130; T41432; T41192; S62524

R;Lye, G.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.

submitted to the EMBL Data Library, November 1995

A;Reference number: Z21830

A;Accession: T39130

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-1374 <LYE>

A;Cross-references: EMBL:AL032824; PIDN: CAB37423.1; GSPDB:GN00068; SPDB:SPCC584.10C

A;Experimental source: strain 972h-; cosmid C584

R;Lye, M.; Rajandream, M.A.; Barrell, B.G.; Seeger, K.; Harris, D.

submitted to the EMBL Data Library, April 1999

A;Reference number: Z21976

A;Accession: T41192

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 722-1374 <LYN>

A;Cross-references: EMBL:AL049662; PIDN: CAB41233.1; GSPDB:GN00068; SPDB:SPCC188.13C

C;Genetics: <LYE1>

A;Gene: SPAC8A4.08c

A;Map position: 1

C;Genetics: <SEE1>

A;Map position: 3

F;32-39/Region: nucleotide-binding motif A (P-loop) #status atypical

F;141-146/Region: nucleotide-binding motif B

F;145-148/Region: DEAD/H motif #status atypical

Query Match 7.6%; Score 402.5; DB 2; Length 1374;

Best Local Similarity 25.6%; Pred. No. 8.8e-13;

Matches 175; Conservative 107; Mismatches 244; Indels 157; Gaps 32;

Db 304 PELQRLPYQMEVAQPALEGKRNIIICLPTGSKTRVAYYIYAKDHDLLKK-----KKAA 353

9 PQL-LRKYQQDVYNIASK-QNTILYVMRTGAKTLLAVKLIIQKLEQILIQQESNLLEHKKI 66

Query Match 7.4%; Score 392.5; DB 2; Length 784;

Best Local Similarity 22.9%; Pred. No. 1.3e-12;

Matches 169; Conservative 128; Mismatches 235; Indels 205; Gaps 28;

Db 306 LQRRPYQMEVAQPALEGKRNIIICLPTGSKTRVAYYIYAKDHDLLKKCASEPGKVVVLVNK 365

19 LEARLYQQIIIAANALKKTLCV-LSTGLGKTAIALVITAGLILTKK-----DGKVLLLAPS 72

Query Match 7.4%; Score 392.5; DB 2; Length 784;

Best Local Similarity 22.9%; Pred. No. 1.3e-12;

Matches 169; Conservative 128; Mismatches 235; Indels 205; Gaps 28;

Db 366 VLLVEQLFRKEFQPFLKKWYRIGLSDGTQKISFPEVVKSCDII 410

67 S-----VELVNKVPLVFCQAEYIYRSOLPAKVGMFYGEULSIEMSEQL--LTNIIILKVNVI 118

Query Match 7.4%; Score 392.5; DB 2; Length 784;

Best Local Similarity 22.9%; Pred. No. 1.3e-12;

Matches 169; Conservative 128; Mismatches 235; Indels 205; Gaps 28;

Db 411 ISTAQILENSLLNLENGEDAGVQLSDFSLIIIDECCHTNKEAVNNIMRHYLMQKLKNRR 470

119 VITADLF--YLFLARG--FLSINDNLII FDECHHAIGNDAYARIMNDY-----HR 166

Db 471 LKKE-NKPVIPVPLPQILGLTASPVGVGGATKQAKAEEHILKLCANLDAFTLTKVKENLDQLK 529

167 AKAVLSKQHFTLPRFGMTASP-----FTGK--KGNLYHRL 200

Query Match 7.4%; Score 392.5; DB 2; Length 784;

Best Local Similarity 22.9%; Pred. No. 1.3e-12;

Matches 169; Conservative 128; Mismatches 235; Indels 205; Gaps 28;

Db 530 NQIQQEPCKKFAIADATREDPFKEKLEIMTRIQTYCQMSPMSPMSDFGTQPYEQWAIQMEKKA 589

201 YQ-----WEQLFDSKAHV--VSNEELADYFCLPEESEYMSNKLV 238

Query Match 7.4%; Score 392.5; DB 2; Length 784;

Best Local Similarity 22.9%; Pred. No. 1.3e-12;

Matches 169; Conservative 128; Mismatches 235; Indels 205; Gaps 28;

Db 590 AKKGNRKERVCAEHL----PKYNEALQINDTIRM----IDAYTHLETFYNEEKDKK 637

239 VPPSDSIIKKCEETLQGCKLJISRAVKTALA--ETIDMGLWFGEQWVLYLVDFETVKRLKK 296

Query Match 7.4%; Score 392.5; DB 2; Length 784;

Best Local Similarity 22.9%; Pred. No. 1.3e-12;

Matches 169; Conservative 128; Mismatches 235; Indels 205; Gaps 28;

Db 638 FAVIEDDSDEGGDDEYCDGDEDDEDDLKKPLKLDETDRFLMLTFFENNKMLKRLAENPEYE 697

Db	220	PNEFKRALK--LINEALKERKLILKDGAVINSIADVTKTTELIELNNRKLFSY-----	268	QY	476	KPVIPPJPOIILGLTASPVGCGATKQAAEHHILKLCANLDAFTIKTKENLIDQQLKQIQEP	535
QY	574	GTQPYEQWAIQMEKKAAKKGKRNKERVCAEHLRKYNEALQINDTIRMIDA-----YTHLE	627	Db	171	KPV--PRVGLGLTASL-IKAKVAPEKULMEQLKKLESAMDS-VIETASD-LVLSLKYGAKP	224
Db	269	------DEEVKRYELIKVCSEALKLMHAKELLESQGKSVFNYIN	306	QY	536	------CKKFAIA--DATTEDPFKEKELIMTRIQTYCQMSPMSPMSDFGTQ-----	576
QY	628	TFYNEEKDKKFAVIEDDSDEGGDDEYCDGDEDEDLKKPLKLDETDRFLMTLFENNKML	687	Db	225	YEVVIIKKDFEIGCLGIPNFDFTVIEIFDETVAFVNNTTEFHFDLDDPDKDSLKTTR	284
Db	307	KLSMORTKSAKSIVNDE-----KVREAVNLIMKSDVEHPKL-	342	QY	577	------PYEQWAIQMEKKAAKKGKRNKERVCAEHLRKYNEALQIND-TIRMID-AYTHLE	627
Qy	688	KRLAENPEYENKEKLTKLRNTIMEQYTRTEESARGLIIFTKTRQSYAL-----SQWITE	747	Db	285	AVFRLQGPWAAM-----RTAQVWKEKELGKJIKSQVLPDKTLRFLNMAKTSMI	331
Db	343	------GKVVDMVKNILEK-----NKDERIIIFQAQYDVTVEKIVNLTON-----	381	QY	628	TFYNEEKDKKPAVIEDDSDEGGDDEYCDGDEDEDLKKPLKLDETDRFLMTLFENNKML	687
Qy	748	GVKAHHLIGAGHSSEFKPMTQNEQKEVISKFRGKINLIAATTVAEGLDIKECNIVIRY	807	Db	332	TI-----KRLI-----EPEMRKLIKSIEARLPYVPQRVIRLFIEL	365
Db	382	GIKAIRFIGQA-NKEKGKMSQKEQIAIERFKK-EGSVLVSTSVSEGIDIPSVNVIIFY	439	QY	688	KRLAENPEYENKEKLTKLRNTIMEQYTRTEESARGLIIFTKTRQSYAL-----SQWITE	740
Qy	808	GLVTNEIAMIWQARGRRADE--STYVLAHSGSGVIEHETVNDFREKMYKAIHCVQNM-	864	Db	366	ETF--NPEFQKERM-KLEKA-----EHLSSAIIFVFDQRYTAYSLIMMRHTKSW---	410
Db	440	EPVPSSEIRFIQRGRAMRGEGGKVVLIAKGTADEAYRSAL-YKEREMKRLL--KNCMC	495	QY	741	NEKFAEVGVKKAHHILAG----HSSEFKPMTQNEQKEVISKFRGKINLIAATTVAE EGL	796
Qy	865	------KPEEYAHKTILELQMQSIMEKKMKTKRNIAKHYK-----NNPSLTLFLCKN	909	Db	411	EPFKF--FVNPDYVVGASGRNLASSDSQGL-HKRCQTEVLRFRFRNEINCLIASTS VLEEGV	467
Db	496	YLNKRLQKKFEKSKKEIKEKEIESTKTAKEETKEEKKTPVTLDFIKQ	555	QY	797	DIKECNIVTRYGLVTNEIAIMVQARGRADESTYVLAHSG-----SGVIE	842
Qy	910	CSV--LACSGED-----INVIEKMHVHNMTP-----EFFKL---YIVRE	943	Db	468	DVKQCNLIVKDFPDRPLDMRSYVQSKGRARRAGSRVYITVEEKDTAAYC SKLPSDIFTLVP	527
Db	556	IEVKERSKSEEDKIKQEIKIPKPKIIVDVRKNNMAKLLHNYANIELKITLEVGDYVLSD	615	QY	843	HETVNDFREKMYKAIHCVQNMKP--EYEAHKILLELQMOSTIMEKKMKTTRNNTA KHYKNNP	900
Qy	944	NKALQKKCADCYQINGEI	960	Db	528	HNQIPIEENGWTK--YCAELLPLPINSPIKHAI-----VLTKNPMPNPKT-----	569
Db	616	RVVVERKTAEDFVNNSII	632	QY	901	SLITFLCKNCVSYLACSGE-DIHVIETK-----MHHVNMTP-----FKE	937
				Db	570	AQMAVALEACRQLHGELEDDNLIPKGRESIAKLEHIDEEPEDEYAPGIAAKVGSSKRKQ	629
				QY	938	LYIVRENKAL-----QKCKADY	954
				Db	630	LYDKRKTARALNESFVEADKECFY	653
				RESULT	12		
				E69055			
				C;Species: Methanobacterium thermoautotrophicum (strain			
				C;Date: 10-Sep-1999 #sequence_change 10-May-2001			
				C;Accession: E69055			
				R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; Jiwani, N.			
				Jiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Reeve, J.N.			
				J.; Bacteriol. 179, 7135-7155, 1997			
				A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: Funct			
				A;Reference number: A69000; PMID:98037514; PMID:9371463			
				A;Status: preliminary; nucleic acid sequence not shown; translation not shown			
				A;Molecule type: DNA			
				A;Residues: 1-738 <MTH>			
				A;Cross-references: GB:AE000903; PID:92622514; PID:92622521			
				A;Experimental source: strain Delta H			
				C;Genetics:			
				A;Gene: MTH1415			
				C;Superfamily: ATP-dependent RNA helicase eIF-4A			
				C;Keywords: ATP; nucleotide binding; P-loop			
				F;38-45/Region: nucleotide-binding motif A (P-loop)			
				F;134-139/Region: nucleotide-binding motif B			
				F;138-141/Region: DEXH motif			
				Query Match	6.6%	Score 350; DB 2; Length 1822;	
				Best Local Similarity	23.5%	Pred. No. 6.5e-10;	
				Matches 175; Conservative	118; Mismatches 249; Indels 202; Gaps 37;		
Qy	309	RPYQMEVAQPALLEGKNNIIICLPTGSGKTRVAVYIAKDH-----LDKKKKKASSEPGKVV	361	Db	14	RDYQVELLDKATK-KNTIVQLGTGSGKTFIAVLLKEYGVQLFAPLDQGGK----RAFF	67
Db	362	LVNKVLLVQEQLERKEFQFLKKWYRVRIGLSGDT-----QLKISFPEVVKSCDIIISTAQI	416	Qy	68	VVEKVNVLVEQ--QAIHIEVHTSFKVQVHGQTSSGLWDSKEQCDQFMKRHHVVVITAQC	124
Qy	417	LENSLLNLENGEDAGVQLSDFSLIIDECHHT-NKEAVNNIMRHYLMOQLKNNRLKKEN	475	Db	125	LDLIRHAYLRIEDMCVLIIFDECHHALGSQHPYRSIMVDY-----KLLRD	170
Db				Query Match	6.4%	Score 342; DB 1; Length 738;	
				Best Local Similarity	24.7%	Pred. No. 4.6e-10;	
				Matches 172; Conservative	106; Mismatches 234; Indels 184; Gaps 31;		
				Qy	302	PEPELQLRPIQMEVAQPALLEGKNNIIICLPTGSGKTRVAVYIAKDHLDKKKKASSEPGKVV	361

12 PE-KIEARTYQQLLAADVIRKGNSMIVAPTALEGKTVVAVLVAERLRKYRGS---KVL 66  
 362 L-VNKVLLVEQLFRKEFQPFLLKWRVIGLGSQDTQLKISF-PEVK---SCDILLISTAQ 415  
 67 LSPSKPLAIQH-EESREFM---IATCTSLTGSIKPEERKERWIKSQIIISATPQ 116  
 416 ILENSLNLLENEDAG-VQLSDPFSLIIIDECHHTNKEAVYNNIMRHYLMQKLKNNRLKKE 474  
 117 TVESDIL---AGRYDLRQDVSLLIVEDECHRAVGTSYVFLASNY---IQNAR--- 161  
 475 NKPVIPLPQIQLGLTASPGVGGATKQAKAEEHTLKCANL--DAFTIKT---VKENLDQ 527  
 162 ---HPLLIGLTTASPGAD---EDKIKTVCENLFMNEVVVKTEGDPVRPVLKP 207  
 528 LKNOIQEPCKKFAIAADATREDPFKEKLLIMTRIQTYCQMSPMNSDFGTOQYEAQIMEK 587  
 208 IKIEWVK---VRMTPELEDIRELLRKVLKN-RLKMLKNLGVID---TISVGK 252  
 588 KAAKKG---NRKERVCAEHRLRKYNEALQINDTIRMIDAYTHLETFYNEEKDKKFAVIE 642  
 253 KDLLKARGRVQNRSTARSTSPPRACYRAISLLASCINVEHALELLET--- 298  
 643 DDSDEGGDDEYCDGDEDDEDDIKPKLQKLDETDRFLMLLFFENNMKLRKLAENPEY---  
 299 ---QG---IRPL---HQYLRLKEKTTKAAGKLLADPDTFRAMLH 334  
 697 ---ENEKLTKLRNTIMEQYTRTBEESARGIIIFTKTRQSAIALSQWITENEKFAEV 747  
 335 TRRAMMSGVEHPKLDRLME-TLKRRELKGDE-ARIIVFTQFRDTLEIYQ---RCKRE 386  
 748 GVKAHHLIGAGHSSEFKPMTQNEQKEVISFKERTGKINLLIATTVAEGLDIKECNIVRY 807  
 387 GINAVKFTYQNSRSRGKSLQKQORDIJKSFRMGNHDVLLSTSVAEEGIDIPSVDLVVMY 446  
 808 GLVTNBELAMQARGRARADESTIVLVAHSGSGVIEHETVNDFREKMYKATHCVQNMKPE 867  
 447 EPVSEIRMIQRRGRTGRKRGKRMVY---LITEKTRD---EAYYYSSIRKERSMK-E 496  
 868 EYAHKILLELQMSIEMEKKMKTKRNIAKHYKNNPSLITFLCKNCVSLACSGEDIHVIEKMH 927  
 497 NLRGGSVNVEVNPIMEP---SGEGPFYIADSR 525  
 928 HVN-MTPEFKELYIIVRENKAQKRCADYQINGEII 961  
 526 EVNSRVLRELKKIGVDFELKPL--AVGDYQISEDITI 559

RESULT 1.3  
 DB4386 ATP-dependent RNA helicase homolog eIF-4A [Imported] - Halobacterium sp. NRC-1  
 C;Species: Halobacterium sp. NRC-1  
 C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 10-May-2001  
 C;Accession: D84386  
 R;NG, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablon, J.; Jung, K.H.; Alam, M.; Freitas, T. Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
 A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li, A;Title: Genome sequence of Halobacterium species NRC-1.  
 A;Reference number: A84160; MUID:20504483; PMID:11016950  
 A;Accession: D84386  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-784 <STO>  
 A;Cross-references: GB:AE004437; NID:910581766; PIDN:AAG20456.1; GSPDB:GN00138  
 C;Genetics:  
 A;Gene: eif4a  
 C;Superfamily: ATP-dependent RNA helicase eIF-4A

Query Match 5.9%; Score 313.5; DB 2; Length 784;  
 Best Local Similarity 21.6%; Pred. No. 1.5e-08;  
 Matches 126; Conservative 91; Mismatches 190; Index 175; Gaps 22;

RESULT 14  
 S48436 probable RNA helicase YIR002C - yeast (Saccharomyces cerevisiae)  
 N;Alternate names: hypothetical protein YTB2C  
 C;Species: Saccharomyces cerevisiae  
 C;Date: 02-Dec-1994 #sequence\_revision 02-Dec-1994 #text\_change 19-Apr-2002  
 C;Accession: S48436; S50885  
 R;Badcock, K.; Churcher, C.  
 Submitted to the EMBL Data Library, August 1994  
 A;Reference number: S48432  
 A;Molecule type: DNA  
 A;Residues: 1-993 <BAD>  
 A;Cross-references: GB:Z47047; EMBL:Z38062; NID:9603997; PID:9763347; MIPS:YIR002C  
 R;Voss, H.; Tamames, J.; Teodoro, C.; Valencio, A.; Sensem, C.; Schwager, C.  
 Yeast 11, 61-78, 1995  
 A;Title: Nucleotide sequence and analysis of the centromeric region of yeast chromosome 1.  
 A;Reference number: S50795; MUID:95282515; PMID:7762303  
 A;Accession: S50885  
 A;Status: nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-993 <VOS>  
 A;Cross-references: EMBL:X79743  
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1994  
 C;Genetics:  
 A;Gene: SGD:MPH1  
 A;Cross-references: SGD:S0001441  
 A;Map position: 9R  
 C;Keywords: ATP; nucleotide binding; P-loop  
 F;107-114/Region: nucleotide-binding motif A (P-loop)  
 F;205-210/Region: nucleotide-binding motif B  
 F;209-212/Region: DEAD/H motif

Query Match	5.7%; Score 303.5; DB 2; Length 993;	Qy	303 EP-ELQLRPyOMEVAQPALEGKNIICLPTGSQKTRAVVYIARDHLDKKKKASEPGKVVIV 361
Best Local Similarity	20.3%; Pred. No. 6.7e-08;	Db	12 EPGAVERRDYQVGLAEQAIR-ENCIVLPTGLGKTAVALQVIAHYLDEGRGALFLAPTRV 70
Matches	143; Conservative 118; Mismatches 246; Indels 197; Gaps 25;	Qy	307 QURPYOMEVAQPALEGKNIICLPTGSQKTRAVVYIARDHLDKKKKASEPGKVVIV 366
Db	86 EVRDYQYTTIVHKSL-FQNTLCAIPTGMGKTFIATSTVMLNYFRWTCKA---KLIIFTAPTR 140	Db	362 LVNKVLLVEQLFRKEFQPFLLKWKYRVTIGLSDTQL--KISFPEVVKSC--DIIISTAQIL 417
Qy	367 LVVEQLFRKEFQPFLLKWKYRVLIGLSDTQL--TOLKISFPEVVKSCDIIISTAQILENS 420	Db	71 LVN------OHRQFLG--RAITISDITLVTDGETIPRRKKAWGGSVICATPEIA 116
Db	141 ELVAQQ----IKACLGITGIPSDQTAIILDKSRKRNREETWANKRVEFFATPQVVEND 192	Qy	418 ENSLNLNLENGEDAGVQLSDFSLIIIDECHHTNKEAVYNNIMRHYLMQKLKNRNLKKKENKP 477
Qy	421 LINLENGEDAGVQLSDFSLIIIDECHHTNKEAVYNNIMRHYLMQKLKNRNLKKKENKPVIP 480	Db	117 RNDI-----ERGLVPLEQFGLVIFDEAHRAVGDIAYSSIA 153
Db	193 L-----KRGVLDPKDTIVCLVIDEAHRATGSSAYTNVVKF-----IDRFNNSY-- 234	Qy	478 VIPLQIQLGLTASPVGGGATKQAKAEEHILKLCANLDLQKNIQEPCK 537
Qy	481 LPQIQLGLTASPVGGGATKQAKAEEHILKLCANLDLQKNIQEPCKKF 539	Db	154 VGDNSRMVGMATL----FSEREKADE----IMGTLLRSRIAQRTEDDPDVKPVQVETAT 205
Db	235 -RLLALTATP-----ASDLEGQEVVNLDISKIEIRTEESMDIVKYMKTKRKKEKI 284	Qy	538 KFAIADATREDPFKEKLIIMTRIQTYCQMSPMS-DEGTQPEQWAIQMEKKAAKGKGNRK 596
Qy	540 AIAADATREDPFKEKLIIMTRIQTYCQMSPMSDFGTQPEQWAIQM-----EKKAA 590	Db	206 EWIKVDLPPPEMKEIQRLLKALD-ERYSSLKRCGYDLGSNRLSALLRL-RMVVLGGNRR 263
Db	285 EV-----PLLIETDIE-----QLGMAVKPVLOQAAIELGIYEECDPSQINAF 327	Qy	597 ERVCAEHLRKYNEALQINDTIRMIDAY--THLETFYNEEKDKKFAVIEEDDSDEGGDDEYC 654
Qy	591 KKGNNRKERVCA----EHLRKYNEAL----QINDTIRMIDAYTHLETFYNEEKDKKFAV 640	Db	264 A-----AKPLFTAIRITYALNIFEAHGVTPLKFCERTSKKGKVVAE----- 306
Db	328 KAMQQSQKIIANPTIFEGIKWNRNFFILQILNNVGQMLKLRKLY-GIRTFFNYYFQNK---	Qy	655 DGDEDDEDLKKPLKLDETDRFLMLTFLFENNNKMLKRLAENPEYENENKLTKLRNTIMEQYTR 714
Qy	641 IEDDSDEGGDDEYDDEDDLKKPLKLDETDRFLMLTFFEN-NKMLKRLAEN---PE 695	Db	307 -----LFEQDR-----NFTGAIARAKAAQAAAGMEHPKIPKLEDAV----- 341
Db	383 -----CTEFTTKYLNKK-----STNKIAAEFFYHPILKNTKRNQCEYNLSDK 424	Qy	715 TEESARG--IIFTKTROSAYALSQWITENEKFAEVGVKAHHLL-AGHSSEEFKPMTQNEQ 771
Qy	696 YENE-KLTKLRLRTIMEQYTRTRTEESARGLIFTKTROSAYALSQWI-----	Db	342 --RGARGKALVFTSYRDSVDLI-----HSRLKAAGTNSGILIGKAGE---KGLKQRKQ 389
Db	425 FVGHGKLQCVRDELMDFFQKRGSDSRVIFTTELRESALEIVKFIDSYADDQIRPHIFIGQ 484	Qy	772 KEVISKPRTGKINLLIATTVAEEGLDIKECNIIVRYGLVTNEIAMVQAGR-ARADESTY 830
Qy	739 -----TENEKFAEVG-----VKAHHLLIGAGH 759	Db	390 VETVAKERDGGDVLSRVGEGLDISEVNVLIVFYDNVPSIIRYVQRGRITGRKDAGRL 449
Db	485 ARAKEGFDEVKYTRKHAPKGRKKVERLHRQEZEKFLEARTKRAANDKLERSARRTGSE 544	Qy	831 VLVAHSGS 838
Qy	760 SSEFKPMTQNEQKEVISKFKRTGKINLLIATTVAEEGLDIKECNIIVRYGLVTNEIAMVQA 819	Db	450 IVLMAKGT 457
Db	545 EAQISGMNQKMQKEVHNFKGEYNVLVCTSTIGEEGLDIGEVDLICYDTTSSPIKNIQR 604	Search completed: June 2, 2004, 19:12:27 Job time : 31 secs	
Qy	820 RGR-ARADESTYVLVHSGSGVIEHETVNDFREKOMYKAIIHCVNMK-----PE-	Db	
Db	605 MGRITGRKRDGKIVLLEFSSNESYKFERAMEDYSTLQALISKQCIDYKKSDRIIPEDIPEC 664	Qy	
Qy	868 -----EYAHKILELQMQS1IMEKKMKTKRNIAK 894	Db	
Db	665 HETLITINDELINEMEDDVDEVIRYATQCMGMGKKVKPKKKAITK 708	Qy	

## RESULT 15

T31323 ATP-dependent RNA helicase, eIF-4A family - Cenarchaeum symbiosum

C;Species: Cenarchaeum symbiosum  
C;Accession: T31323  
C;Sequence: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 11-Jan-2000R;Schleper, C.; DeLong, E.F.; Preston, C.M.; Feldman, R.A.; Wu, K.Y.; Swanson, R.V.  
J. Bacteriol. 180, 5003-5009, 1998A;Title: Genomic analysis reveals chromosomal variation in natural populations of the un  
A;Reference number: Z20994; MUID:98422450; PMID:9748430  
A;Accession: T31323  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-502 <SCH>

A;Cross-references: EMBL:AF083072; NID:93599393; PID:93599409; PIDN: AAC62714.1

Query Match 5.3%; Score 284; DB 2; Length 502;  
Best Local Similarity 24.3%; Pred. No. 2.6e-07;  
Matches 133; Conservative 94; Mismatches 207; Indels 114; Gaps 23;



-!- FUNCTION: Probably involved in cleaving double-stranded RNA in the RNA silencing pathway. It produces 20 to 23 bp dsRNAs (siRNAs and miRNAs) which target the selective destruction of homologous RNAs.

-!- SUBUNIT: May interact with ARGONAUTE1 or PINHEAD through their common PAZ domains (By similarity).

-!- SUBCELLULAR LOCATION: Nuclear (Potential).

-!- TISSUE SPECIFICITY: Highly expressed in flowers and seeds and detected in leaves and stems. Found in ovule integuments, inflorescence and floral meristems, stigma of flowers until just before pollination, vasculature of the funiculus, and embryo.

-!- DEVELOPMENTAL STAGE: Detected in the embryo, but not in the suspensor, up to the globular stage.

-!- MISCELLANEOUS: Expression in the early embryo is from the maternally contributed genome.

-!- SIMILARITY: Belongs to the helicase family.

-!- SIMILARITY: Contains 2 RNase III domains.

-!- SIMILARITY: Contains 1 PAZ domain.

-!- SIMILARITY: Contains 2 DRBM (double-stranded RNA-binding) domains.

-!- CAUTION: Ref.3 sequence differs from that shown due to erroneous gene model prediction.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; AF187317; AAF03534.1; DR EMBL; AF292941; AAG38020.1; DR EMBL; AF292940; AAG38019.1; DR EMBL; AC007323; AAF26461.1; ALT\_SEQ. DR InterPro; IPR001410; DEAD. DR InterPro; IPR001159; DS\_RBD. DR InterPro; IPR005034; DUF283. DR InterPro; IPR001650; Helicase\_C. DR InterPro; IPR003100; PAZ. DR InterPro; IPR000999; RNase\_III. DR Pfam; PF00270; DEAD; 1. DR Pfam; PF00035; dsrm; 2. DR Pfam; PF03368; DUF283; 1. DR Pfam; PF00271; helicase\_C; 1. DR Pfam; PF02170; PAZ; 1. DR SMART; SM00636; Ribonuclease\_3; 2. DR SMART; SM00487; DEXDC; 1. DR SMART; SM00358; DSRM; 2. DR SMART; SM00490; HELICC; 1. DR PROSITE; PS50137; RIBOC; 2. DR PROSITE; PS50137; DS\_RBD; 2. DR PROSITE; PS50821; PAZ; 1. DR PROSITE; PS00517; RNASE\_3\_1; 1. DR PROSITE; PS50142; RNASE\_3\_2; 2. KW Hydrolase; Nuclease; Endonuclease; Helicase; Nuclear protein; ATP-binding; RNA-binding; Repeat.

NP BIND 249 256 ATP (POTENTIAL). DOMAIN 1180 1318 PAZ. FT DOMAIN 1342 1518 RNASE\_III\_1. FT DOMAIN 1559 1707 RNASE\_III\_2. FT DOMAIN 1733 1796 DRBM\_1. FT DOMAIN 1831 1906 DRBM\_2. FT SITE 378 381 DECH\_BOX. FT MUTAGEN 415 415 P->S: IN SIN1-1; IMPAIRED REPRODUCTIVE DEVELOPMENT. FT MUTAGEN 431 431 I->K: IN SIN1-2; IMPAIRED REPRODUCTIVE DEVELOPMENT.

NDICLRK->IAEIDPG: IN CAF-1; CONVERTS THE FLORAL MERISTEM TO AN INDETERMINATE STATE. MISSING: IN CAF-1; CONVERTS THE FLORAL MERISTEM TO AN INDETERMINATE STATE. S -> F (IN REF. 1 AND 2).

CC	FT	CONFLICT	988	988	Y -> H (IN REF. 1 AND 2).
CC	SQ	SEQUENCE	1909	AA;	213572 MW; EAA944F0C0C81D4C CRC64;
CC	Query	Match	7.6%	Score 406;	DB 1; Length 1909;
CC	Db	Best Local Similarity	23.0%	Pred. No. 2.2e-12;	
CC	Db	Matches 199; Conservative	131;	Mismatches 271; Indels 264; Gaps 40;	
CC	QY	135 LDKCME----BELLTIEDNRNRIAAENNGNEGSYRBLKRRIVQENWESAFLNVLQRT 188			
CC	Db	18 LDACEDISCDLIDDDLVSEDFSSVAVNEST-DENGV-----INDEFFGGIDHILDSI 67			
CC	QY	189 GNNEVQELTGSDCSESNAEIENLSQLQVDPQVEEQLLSTTVQPN-LEKEVWGMENNSSES 247			
CC	Db	68 KNG---GGLPNNGVSDTNSQINEVTT-PQV---IAKETVKENGQKNGKRDDEFSEE 119			
CC	QY	248 SFAD-----SSVVSSESDTSLAEGSVSCLDE-----SLGNH-----277			
CC	Db	120 GDKDRKARVCSSYQSERNSLNSGRHTVNSREGDRFMNMRKTRNWEAGNNKKCRECNYR 179			
CC	QY	278 -----SNTMGSDD-----BENVAARAS 301			
CC	Db	180 RDGRDREVRYWERDKVGSNNELVYRSGTWEADHERDVKKVSGNREC DVKAEN---KSK 236			
CC	QY	302 PEP-----ELOLRPYOMEVAQPALEGKNIILICLPTGSGKTRVAVY---IAKDHLDKKK 351			
CC	Db	237 PEERKEKVVEEQAQARRYQDVLQEA-KAKNTIAFLLETGAGKTLIAILLIKSVHKDLMSONR 295			
CC	QY	352 KASEPGBKVIVLYNKVLLVEQ---LFFKEF-----QPF---LKKWYRVIGLSGDT 394			
CC	Db	296 KMLs---VFLVPKVPLVYQQAEVIRNQTCFQVGHYCGERGQDFWDSRWRQREF-----345			
CC	QY	395 QLKISFPEVVKVSCDILISTAQILENSLNLQENGEDAGVQLSDFSLIILDECHTNTKEAVY 454			
CC	Db	346 -----ESKQVLVMTAQILLNIL-----RHSIIRMETIDLILDECHHAVKKHPY 389			
CC	QY	455 NNIMRHYLQKLLKRNRLKENKPKVILPQLLGTASP-GVGGATKQAKAEEHLKLCAAL 513			
CC	Db	390 SLVMSSEFYHTTPKDKR-----PAIFGMMTASPVNLKGVSSQVDCAIKIRNLLETKL 438			
CC	QY	514 DAFTIKTKVENLDQLKQIQLQEMKKAQKGNRKERVCAEHLRKYNEALQINDTIRMID-----A 622			
CC	Db	439 DS-TVCTIKDR-KELEKHTVPMMPSETIVYEDKAATTMWSLHETITKOMIAAEEAAQASSRK- 495			
CC	QY	573 FGTOPQEWAQWALQMEKKAQKGNRKERVCAEHLRKYNEALQINDTIRMID-----A 622			
CC	Db	496 -----SKWQFMGARDAGAKDELRQWYGVSETS ESDGAANTLHKLRAINNTLAELGOWCA 549			
CC	QY	623 YTHLETF-----YNEEKDKKF-----AVIE-----642			
CC	Db	550 YKYGOSFLSALQSDERVNQFDVKFQESYLSLEVSLQCELLLEGAAEKVAAEVGKPGENG 609			
CC	QY	643 ---DDSDEG--GDDEYCDGDEDIDLKKPLKLDDETDRFLMTLFFENNKKMLKRLAENPEYE 697			
CC	Db	610 NAHDEEEGELPDDPPVSGGEHVDVIGAVAD-----GKVTPKV-----649			
CC	QY	698 NEKLTKLNRNTIMEQYTRTEESARGLFTKTKTROSAYALSQWITENEKFAEVG----VKAHH 753			
CC	Db	650 -QSLIKL---LLKQHTAD-FRAIVEVERVVAALVLP-----KVEAELPSLSFIRCAS 697			
CC	QY	754 LIGAGHSSEFKPNTQNEQKEVISKFRTGKINLLIATTVAEEGLDIKECNIVTRYGLVTNE 813			
CC	Db	698 MIGHNNSQEMK---SSQWQMDTISKFRDGHTVLLVATSAEGLDIRQCNVVMRFDIAKTV 754			
CC	QY	814 IAWVOARGRARADESTYVLAHSGS 838			
CC	Db	755 LAYIQSRGRARKKGPGSDYIILMVERGN 779			

RESULT 2  
YQ3D SCHPO  
ID YQ3D SCHPO STANDARD;  
AC Q09884; Q9UN1;  
DT 01-FEB-1996 (Rel. 33, Created)

DT	SQ	SEQUENCE	1374 AA;	158039 MW;	89AE9EF8DE7966C6 CRC64;
DT 01-FEB-1996 (Rel. 33, Last sequence update)	Query Match	7.6%;	Score 402.5;	DB 1;	Length 1374;
DT 28-FEB-2003 (Rel. 41, Last annotation update)	Best Local Similarity	25.6%;	Pred. No. 2.1e-12;		
DE Putative helicase C188.13c (EC 3.6.1.-).	Matches	175;	Conservative	107;	Mismatches 244; Indels 157; Gaps 32;
GN SPCCI188.13C OR SPCCC584.10C.					
OS Schizosaccharomyces pombe (Fission yeast).					
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomyces; Schizosaccharomyces pombe.					
OC Schizosaccharomyces pombe.					
OC Schizosaccharomyces pombe.					
OX NCBI_TaxID=4896;					
RN SEQUENCE FROM N.A.					
RC STRAIN=972;					
RX PMID=21848401; PubMed=11859360;					
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A., Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S., Brooks K., Brown S., Chillingworth T., Churcher C.M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A., Collins M., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G., Gentles S., Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K., James K., Jones L., Jones M., Leather S., McDonald S., McLean J., Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C., Oliver K., O'Neil S., Pearson D., Quail M.A., Rabbinowitsch E., Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S., Skelton J., Simmonds M., Squares R., Squares S., Stevens K., Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S., Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B., Weltjens I., Vansstreels E., Rieger M., Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H., Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M., Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B., Goffeau A., Cadieu E., Dreano S., Gloux S., Lelauvre V., Mottier S., Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M., Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G., Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J., Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L., Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J., Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.; RT "The genome sequence of Schizosaccharomyces pombe.";					
RL Nature 415:871-880(2002).					
CC -!- SIMILARITY: Belongs to the helicase family.					
CC -!- SIMILARITY: Contains 2 RNase III domains.					
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).					
CC EMBL; AL032824; CAB37423.1; -.					
DR EMBL; AL049662; CAB41233.1; -.					
DR PIR; T39130; S62524.					
DR GenedB SPombe; SPCCC1BB.13.C; -.					
DR InterPro; IPR001410; DEAD.					
DR InterPro; IPR005034; DUF283.					
DR InterPro; IPR001650; Helicase_C.					
DR InterPro; IPR000999; RNase_III.					
DR Pfam; PF00270; DEAD; 1.					
DR Pfam; PF03368; DUF283; 1.					
DR Pfam; PF00271; helicase_C; 1.					
DR Pfam; PF00636; Ribonuclease_3; 2.					
DR SMART; SM00487; DEXDC; 1.					
DR SMART; SM00490; HELICC; 1.					
DR SMART; SM00535; RIBOC; 2.					
DR PROSITE; PS00517; RNASE_3-1; 1.					
DR PROSITE; PS50142; RNASE_3-2; 2.					
KW Hypothetical protein; Helicase; ATP-binding; Hydrolase; Nuclease; Endonuclease; Repeat.					
FT NP_BIND 32 39 ATP (POTENTIAL).					
FT SITE 145 148 DECH BOX.					
FT DOMAIN 916 1038 RNASE III 1.					
FT DOMAIN 1083 1233 RNASE III 2.					
RN SEQUENCE FROM N.A.					
RP STRAIN=JAL-1 / DSM 2661 / ATCC 43067;					
RC MEDLINE=96337999; PubMed=8688087;					
RX Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,					

Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D., Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I., Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A., Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D., Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C., Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M., Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;	"Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii"; Science 273:1058-1073 (1996). - SIMILARITY: Belongs to the DEAD box helicase family. DEAH subfamily.	Db	301 KLSMORTKSAKSIVNDE-----KVRAVNLIMKSDVEHPKL- 336
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).		Db	301 KLSMORTKSAKSIVNDE-----KVRAVNLIMKSDVEHPKL- 336
EMBL; MG1505; -.	TIGR; U67591; AAB99518.1; -.	Db	301 KLSMORTKSAKSIVNDE-----KVRAVNLIMKSDVEHPKL- 336
InterPro; IPR001410; DEAD.	InterPro; IPR002464; DEAH box.	Db	301 KLSMORTKSAKSIVNDE-----KVRAVNLIMKSDVEHPKL- 336
InterPro; IPR006166; ERCC4.	InterPro; IPR001650; Helicase_C.	Db	301 KLSMORTKSAKSIVNDE-----KVRAVNLIMKSDVEHPKL- 336
InterPro; IPR00445; HhH.	InterPro; IPR003583; HHH_1.	Db	301 KLSMORTKSAKSIVNDE-----KVRAVNLIMKSDVEHPKL- 336
InterPro; PE00270; DEAD; 1.	Pfam; PF002732; ERCC4; 1.	Db	301 KLSMORTKSAKSIVNDE-----KVRAVNLIMKSDVEHPKL- 336
Pfam; PF00271; helicase_C; 1.	Pfam; PF00633; HHH; 2.	Db	301 KLSMORTKSAKSIVNDE-----KVRAVNLIMKSDVEHPKL- 336
SMART; SM00487; DEXDC; 1.	SMART; SM00490; HELICC; 1.	Db	301 KLSMORTKSAKSIVNDE-----KVRAVNLIMKSDVEHPKL- 336
SMART; SM00278; Hhh1; 1.	SMART; SM00278; Hhh1; 1.	Db	301 KLSMORTKSAKSIVNDE-----KVRAVNLIMKSDVEHPKL- 336
PROSITE; PS00690; DEAH ATP HELICASE; FALSE NEG.	Hypothetical protein; ATP-Binding; RNA-binding; Helicase;	Db	301 KLSMORTKSAKSIVNDE-----KVRAVNLIMKSDVEHPKL- 336
Complete proteome.	Complete proteome.	Db	301 KLSMORTKSAKSIVNDE-----KVRAVNLIMKSDVEHPKL- 336
NP_BIND	NP_BIND	Db	301 KLSMORTKSAKSIVNDE-----KVRAVNLIMKSDVEHPKL- 336
SITE	35 42 ATP (POTENTIAL).	Db	301 KLSMORTKSAKSIVNDE-----KVRAVNLIMKSDVEHPKL- 336
DOMAIN	137 140 DEAH BOX.	Db	301 KLSMORTKSAKSIVNDE-----KVRAVNLIMKSDVEHPKL- 336
SEQUENCE	532 535 POLY-GLU.	Db	301 KLSMORTKSAKSIVNDE-----KVRAVNLIMKSDVEHPKL- 336
	778 AA; 88682 MW; DBC1F220801338D8 CRC64;	Db	301 KLSMORTKSAKSIVNDE-----KVRAVNLIMKSDVEHPKL- 336
Query Match	7.4%; Score 392.5; DB 1; Length 778;	Db	301 KLSMORTKSAKSIVNDE-----KVRAVNLIMKSDVEHPKL- 336
Best Local Similarity	22.9%; Pred. No. 3e-12;	Db	301 KLSMORTKSAKSIVNDE-----KVRAVNLIMKSDVEHPKL- 336
Matches 169; Conservative 128; Mismatches 235; Indels 205; Gaps 28;		Db	301 KLSMORTKSAKSIVNDE-----KVRAVNLIMKSDVEHPKL- 336
Y	306 LQLRPYQMEVAQPALEGQKNTTICLPTGSKTRAVYIAKDHLDKKKKASEPQKVVVLVNK 3.65	Db	301 KLSMORTKSAKSIVNDE-----KVRAVNLIMKSDVEHPKL- 336
	13 LEARLYQQIIAANALKKRTLCV-LSTGLGKTAIIILVIAGILTKK----DGKVLLAPS 66	Db	301 KLSMORTKSAKSIVNDE-----KVRAVNLIMKSDVEHPKL- 336
Y	3.66 VLLVEQLFRRKEFQPFLKWKWYRVLGSGDTQLKLSPEVVVKSCDIIISTAQILLENLLNLE 4.25	Db	301 KLSMORTKSAKSIVNDE-----KVRAVNLIMKSDVEHPKL- 336
	67 RPLVEQHYNRLKQVLNIDEDKIIALTGKQPK-KRAELYKKGKFIATPQVIENDII--- 122	Db	301 KLSMORTKSAKSIVNDE-----KVRAVNLIMKSDVEHPKL- 336
Y	4.26 NGEDAG-VQLSDFSLIIIDECHHTNKEAVNNIMRHYLMQQLKNRNLKKENPKVPLPQI 4.84	Db	301 KLSMORTKSAKSIVNDE-----KVRAVNLIMKSDVEHPKL- 336
	123 ---AGRINDEFILLIADEAHHTTGDHAY----AFVAKKFKD-----CHI 1.61	Db	301 KLSMORTKSAKSIVNDE-----KVRAVNLIMKSDVEHPKL- 336
Y	4.85 LGLTASPGVGATQKAKEEHLKCANLDRAFTIKTKENLDQLK----- 5.29	Db	301 KLSMORTKSAKSIVNDE-----KVRAVNLIMKSDVEHPKL- 336
	162 LGLTASPG-----SDIDKVMCEICENLGTIEVEVRTDEDVVKPYIAYKVKLIPRIDI 21.3	Db	301 KLSMORTKSAKSIVNDE-----KVRAVNLIMKSDVEHPKL- 336
Y	5.30 -NQIQEPCKKFAIADATRE-----DPFKEKELLEIMTRIQTYCQMSPMSSDF 5.73	Db	301 KLSMORTKSAKSIVNDE-----KVRAVNLIMKSDVEHPKL- 336
	214 PNEFKRALK-LINEALKERLKILKDAVGIVNSADVTKTTELIELNNKLFESY----- 26.2	Db	301 KLSMORTKSAKSIVNDE-----KVRAVNLIMKSDVEHPKL- 336
Y	5.74 GTQPYEQWAIQMEKKAAKKGKNRKERVCAEHRLRKYNEALQIINDTIRMIDA-----YTHLE 6.27	Db	301 KLSMORTKSAKSIVNDE-----KVRAVNLIMKSDVEHPKL- 336
	263 -----DEEVKYELIKVCEALKLMAKELLESQGKSFLNLYIN 3.00	Db	301 KLSMORTKSAKSIVNDE-----KVRAVNLIMKSDVEHPKL- 336
Y	6.28 TFYNEEKDKKFAVIEDDSDEGGDDDEDDLKKPLKLDDETDRFLMTLFFENNKKM. 6.87	Db	301 KLSMORTKSAKSIVNDE-----KVRAVNLIMKSDVEHPKL- 336
REVISIONS.		Db	301 KLSMORTKSAKSIVNDE-----KVRAVNLIMKSDVEHPKL- 336
RX	MEDLINE=22158633; PubMed=12168954;	Db	301 KLSMORTKSAKSIVNDE-----KVRAVNLIMKSDVEHPKL- 336
RX	Nakajima D., Okazaki N., Yamakawa H., Kikuno R., Ohara O., Nagase T.,	Db	301 KLSMORTKSAKSIVNDE-----KVRAVNLIMKSDVEHPKL- 336
RT	"Construction of expression-ready cDNA clones for KIAA genes: manual	Db	301 KLSMORTKSAKSIVNDE-----KVRAVNLIMKSDVEHPKL- 336
RT	curation of 330 KIAA cDNA clones".	Db	301 KLSMORTKSAKSIVNDE-----KVRAVNLIMKSDVEHPKL- 336

DNA Res. 9:99-106 (2002).		SEQUENCE OF 1238-1912 FROM N.A.	
SEQUENCE OF 1238-1912 FROM N.A.		TISSUE=Lung;	
MEDLINE=99162526; PubMed=10051563;		Provost P., Samuelson B., Radmark O.; "Interaction of 5'-lipoxygenase with cellular proteins." Proc. Natl. Acad. Sci. U.S.A. 96:1881-1885 (1999).	
-!- FUNCTION: Involved in cleaving double-stranded RNA in the RNA interference (RNAi) pathway. It produces 21 to 23 bp dsRNAs (siRNAs) which target the selective destruction of homologous RNAs.		-!- SIMILARITY: Belongs to the helicase family.	
-!- SIMILARITY: Contains 2 RNase III domains.		-!- SIMILARITY: Contains 1 PAZ domain.	
-!- SIMILARITY: Contains 1 DRBM (double-stranded RNA-binding) domain.		-!- SIMILARITY: Contains 1 DRBM (double-stranded RNA-binding) domain.	
<p>This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).</p>			
EMBL; AB028449; BAA78691.1; ALT_INIT.		QY 309 RPYOMEVAQPALEGKNNIIICLPTGSGKTRVAVYIARD-HLDKKKKASEPGK-VIVLVNKV 366	
EMBL; AB1132261; CAB38857.2; ALT_INIT.		35 RKYQVELLEAZALD-HNTIVCLNTGSGKTFIAVLITKEBSYQIRGDFSRNGKRTVFLVNSA 93	
DR; GO:0005622; C:intracellular; NAS.		Db 367 LLVEQLFREFOFPEFLKWKYRVIGLNSGDTQLK1--SFPEVVKSC--DIII 411	
DR; GO:0003423; F:double-stranded RNA binding; IDA.		94 NQVAQ--QVSAYRTHSDLKVGYESNLNEVNASWTKERWNQEFFXHQVLI 139	
DR; GO:0003423; F:double-stranded RNA binding; IDA.		Db 412 STAQILENSSLNNLENGEDAGVQLSDFSLIILIDECHTNEKEAVYNNIMRHYTLMQKLKNRNL 471	
DR; GO:0003423; F:double-stranded RNA binding; IDA.		140 MTCYVALNVLKN---GYLSLSDINLLVDECHLAILDHPYREIMK- 181	
DR; GO:0003423; F:double-stranded RNA binding; IDA.		Db 472 KKENKPVIPPLPQILGLTASPGVGGATKQAKAEEHTLKLICANLDAFTIKTKYKENLJDQ--- 527	
DR; GO:0003423; F:double-stranded RNA binding; IDA.		182 LICENCP--SCPRLIGLTASI-LNGKCDPPEELEEKIQKL EKILKSNAETATD 230	
DR; GO:0003423; F:double-stranded RNA binding; IDA.		528 -LKNQIQEPCKKFA1AD---ATREDPFKEKLEI---MTRIQTYC 565	
DR; GO:0003423; F:double-stranded RNA binding; IDA.		231 VVLDRTYTSQPCE--IVVDCGPFETDRSGLYERLIMELLEEAALNFINDCNISVHSKERDSTLI 288	
DR; GO:0005622; C:intracellular; NAS.		QY 566 QMSPPMSD----FGTQPYEQWAIQOMEKKAAK- -KGNRKE----RV 599	
DR; GO:0003423; F:double-stranded RNA binding; IDA.		Db 289 SKQILSDCRAVLVVLGPWCAKTYAGMMVRELQKYTKHEQEEELHRKFLLEFTDTFLRKIHAL 348	
DR; GO:0003423; F:double-stranded RNA binding; IDA.		QY 600 CAEHLRKYNN----EALQINDTIRMIDAYTHLE---TFYNEEKDKKFAVIEDDSDE 647	
DR; GO:0003423; F:double-stranded RNA binding; IDA.		Db 349 CEEHFSPPASLDLKFVTPKVKLLEILRKYKPYERQQFESVWYNNRQNQDNYVWSWSDSED- 407	
DR; GO:0003423; F:double-stranded RNA binding; IDA.		QY 648 GGDDEYCDGDEDDELLKKPLKLDETDRLMTFFENNKKMLKRLAENPEYENEKLTKLRLNT 707	
DR; GO:0003423; F:double-stranded RNA binding; IDA.		Db 408 -DDE---DEEEEEKKEKP----ETNFPSPFPTNI 431	
DR; GO:0003423; F:double-stranded RNA binding; IDA.		QY 708 IMEQYTRTRTEESARGLIIFTKTKRQSAYALSQWITENBK----FAEVGVKAHHLIGAGH- - 759	
DR; GO:0003423; F:double-stranded RNA binding; IDA.		Db 432 LC----GIFTVERRYTAVVLNRLIKAEAGKQDPELAYISSNFTIGHGIGKQPER 480	
DR; GO:0003423; F:double-stranded RNA binding; IDA.		Db 760 ---SSEFKPMTQNEQKEVVISFKFRTGKINLLIATTVAEEGLDIKECQNIVRYGLVTNEIA 815	
DR; GO:0003423; F:double-stranded RNA binding; IDA.		Db 481 NKQMEAERF---KQEEVLRKFRAHETNLLIATATIVEGVDIPKCNLYVRFDLPTEYRS 535	
DR; GO:0003423; F:double-stranded RNA binding; IDA.		QY 816 MVQARGRARADESTYVLAHSGGGVIEHETVNDEREKM-MYKAIIHCYQNMK 865	
DR; GO:0003423; F:double-stranded RNA binding; IDA.		Db 536 YVQSKGRARAPISNYIMLA----DTDKIKSFEEDLTKYKAIEKILRNK 579	
<p>RESULT 5</p>			
DICE_MOUSE		ID DICE_MOUSE STANDARD; PRT; 1906 AA.	
Q8R418;		AC 41; Created	
DT 28-FEB-2003 (Rel. 41, Last sequence update)		DT 28-FEB-2003 (Rel. 41, Last annotation update)	
DE Endoribonuclease Dicer (EC 3.1.26.-) (Double-strand-specific ribonuclease MDCR-1).		DT 15-MAR-2004 (Rel. 43, Last annotation update)	
DE ribonuclease MDCR.		GN Mus musculus (Mouse).	
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Murinae; Mus.		OC Mammalia. NCBI_TaxID=10090; NCBI_TaxID=1.	
OX		RN SEQUENCE FROM N.A. PMID=21886641; PubMed=11889553;	
RN		RN Nicholson R.H., Nicholson A.W.;	
RN		RA "Molecular characterization of a mouse DNA encoding Dicer, a ribonuclease III ortholog involved in RNA interference."	
RN		RT Mamm. Genome 13:67-73 (2002).	
RN		RP SEQUENCE FROM N.A.	
SEQUENCE 1912 AA; 217627 MW: 996399DB4B074E21 CB644;		SQ	

RC STRAIN=Czech II;  
 RA Svoboda P., Anger M., Stein P., Schultz R.M.;  
 "Mouse dicer homolog in oocyte and preimplantation embryos.";  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DDJB databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22414943; PubMed=12526743;  
 RA Doi N., Zenno S., Ueda R., Ohki-Hamazaki H., Ui-Tei K., Saigo K.;  
 RT "Short-interfering-RNA-mediated gene silencing in mammalian cells  
 requires Dicer and eIF2C translation initiation factors.";  
 RL Curr. Biol. 13:41-46(2003).  
 -!- FUNCTION: Involved in cleaving double-stranded RNA in the RNA  
 interference (RNAi) pathway. It produces 21 to 23 bp dsRNAs  
 (siRNAs) which target the selective destruction of homologous  
 CC RNAs.  
 CC -!- TISSUE\_SPECIFICITY: Expressed in a wide variety of tissues.  
 CC -!- SIMILARITY: Belongs to the helicase family.  
 CC -!- SIMILARITY: Contains 2 RNase III domains.  
 CC -!- SIMILARITY: Contains 1 PAZ domain.  
 CC -!- SIMILARITY: Contains 1 DRBM (double-stranded RNA-binding) domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@isb-sib.ch).  
 DR EMBL; AF484524; AAL84638.1;  
 DR EMBL; AF430845; AAM21495.1; ALT\_INIT.  
 DR EMBL; ABO81470; BAC15765.1; .  
 DR MGD; MGI:2177178; Dicer1.  
 DR GO; GO:0005622; C:intracellular; ISS.  
 DR GO; GO:003725; F:double-stranded RNA binding; ISS.  
 DR GO; GO:0030423; P:RNA interference, targeting of mRNA for des. . ; ISS.  
 DR InterPro; IPR001410; DEAD.  
 DR InterPro; IPR001159; DS\_RBD.  
 DR InterPro; IPR005034; DUF283.  
 DR InterPro; IPR001650; Helicase\_C.  
 DR InterPro; IPR003100; PAZ.  
 DR InterPro; IPR000999; RNase\_III.  
 DR Pfam; PF00035; derm\_1.  
 DR Pfam; PF003368; DUF283; 1.  
 DR Pfam; PF00271; helicase\_C; 1.  
 DR Pfam; PF002170; PAZ\_1.  
 DR Pfam; PF000336; Ribonuclease\_3; 2.  
 DR SMART; SM000487; DEXDC; 1.  
 DR SMART; SM00358; DSRM; 1.  
 DR SMART; SM00490; HELICC; 1.  
 DR SMART; SM00535; RIBOC; 2.  
 DR PROSITE; PS50137; DS\_RBD; 1.  
 DR PROSITE; PS50821; PAZ; 1.  
 DR PROSITE; PS00517; RNASE\_3\_1; 1.  
 DR PROSITE; PS50142; RNASE\_3\_2; 2.  
 DR Helicase; ATP-binding; RNA-binding; Hydrolase; Nuclease; Endonuclease;  
 KW Repeat.

FT NP\_BIND 34 41 ATP (POTENTIAL).  
 FT SITE 165 168 DECH BOX.  
 FT DOMAIN 881 1032 PAZ.  
 FT DOMAIN 1266 1326 RNASE\_III\_1.  
 FT DOMAIN 1650 1808 RNASE\_III\_2.  
 FT DRBM. DRBM.  
 FT CONFLICT 97 97 A -> C (IN REF. 2).  
 FT CONFLICT 157 157 S -> P (IN REF. 3).  
 FT CONFLICT 279 279 H -> Y (IN REF. 2).  
 FT CONFLICT 600 600 A -> T (IN REF. 3).  
 FT CONFLICT 749 749 E -> D (IN REF. 2).  
 FT CONFLICT 827 827 T -> I (IN REF. 2).  
 FT CONFLICT 878 878 G -> S (IN REF. 1).  
 FT CONFLICT 955 955 Y -> C (IN REF. 3).  
 FT SEQUENCE 1906 AA; 215766 MW; 230EA9BFC19F3091 CRC64;  
 Query Match 6.7%; Score 354; DB 1; Length 1906;  
 Best Local Similarity 24.7%; Pred. No. 8.2e-10;  
 Matches 158; Conservative 92; Mismatches 213; Indels 176; Gaps 29;  
 QY 309 RPYQMEVAQPALEGKNIICLPTGSKTRVAYIAKD--HLDKKKASEPGKTVIVLYNKV 366  
 35 RKYQVELLEAALD-HNTIVCLNTGSKTFIAVLTKELABQIRGDLNPHAKRTFLYNSA 93  
 Db 367 -LIVEQLFRKEFOPFLKKWYRVLGSDTOLKISR----PEVKSCDIIISTAQILEN 419  
 Db 94 NQVAQQVSAVRTHSDLK----VGEYSDLEVNAAWTKERWSOEFTHQ-QVLMITCYV--- 144  
 QY 420 SLLNLENEDAGVQLSDFSLIIDECHTNKEAVYNNIMBHMLMQKLKNNRLKKENKPKVI 479  
 Db 145 ALTVLKNGY--LSSLSDINLLVFDCHLAILDHPYREIMK-----LCESCP-- 187  
 QY 480 PLPQIQLTASPVGYGGATQAKAAEEHILKL--CANLDAFTIKTVKENLQLKNQIQEPCCK 537  
 Db 188 SCPRILGLTASL-LNGKCDPEELLEKIQKLERIERSDAET---ATDLVVLDRYTSQPCE 242  
 QY 538 KFAIAD---ATREDPFPKEKLEI----MTRIQTYCQMSPMNSD---- 572  
 Db 243 -IVVDCGPFIDRSGLYERLLMELAALDFINDCNVAVHSKERDSTLISKQILSDCRAVL 300  
 QY 573 -FGTQPYEQWAIQMEKKAAC--KGNRKE----RVCAEHLR---- 605  
 Db 301 VVLPGPWCADKVAGMVTRELQKYIKHEQEEHLRKFLLFTDLRKHALCEEYFSPASLDL 360  
 QY 606 KY--NEALQINDTIRMDAYTHLE---TFYNEEKDKKFAVIEDDSDEGGDDEYCDGDED 659  
 Db 361 KYVTPKVMKLLIEILRKYKPYERQQFESVIEWYNNRNQDNYVWSDSDEDDDEBIEEKEKP 420  
 QY 660 EDDLKKPLKLDETDRLFLMTFFENNKNMKLKRLNTIMEQYTRTEESA 719  
 Db 421 ETNFPSP----FTN----IL 432  
 QY 720 RGIIFTKTROSAYALSOVITNEK----FAEVGVKAHLLIGAGH----SSEFKPMT 767  
 Db 433 CGIIFVERRYTAVVNLRLIKEAGKQDPELAYISSNFTIGHIGKQNPRSKQMEAEFR--- 489  
 QY 768 QNEQKEVITSKFRTGKINLILATTVAEGGLDIKECNIVYRGVTNEIAMVQARGRARADE 827  
 Db 490 -KQEEVLRKFRAHETNLILIATSVVEGVDIPKCNLVVRFLDPLTEYRSYVQSKGRARAPI 547  
 QY 828 STYVLVAHSGSGVIEHETVNDFREKM-MYKAITHCVQNMK 865  
 Db 548 SNTYMLA----DTDKIKSFEEDLRTYKATEKILRNRK 579  
 RESULT 6  
 DCR1\_CAEEL ID DCR1\_CAEEL STANDARD; PRT; 1845 AA.  
 AC P34529;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Endoribonuclease dcr-1 (EC 3.1.26.-).  
 GN DCR-1 OR K12H4.8.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderrinae; Caenorhabditis.  
 OC NCBI\_TAXID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;

REVISIONS:

RA Waterston R.; Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases. [3]

RA Waterston R.; Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases. [2]

RP REVISIONS.

RA "2.2 Mb of contiguous nucleotide sequence from chromosome III of *C. elegans*." Nature 368:32-38 (1994).

RN [2]

RP CHARACTERIZATION.

RA Ketting R.F.; Fischer S.E.J.; Bernstein E.; Sijen T.; Hannon G.J.; Plasterk R.H.A.; "Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*." Genes Dev. 15:2654-2659 (2001).

CC -1- FUNCTION: Involved in cleaving double-stranded RNA in the RNA interference (RNAi) pathway. It produces 21 to 23 bp dsRNAs (siRNAs) which target the selective destruction of homologous RNAs. Seems to process the precursor of the small temporal RNA let-7 which is involved in developmental timing.

CC -1- SIMILARITY: Belongs to the helicase family.

CC -1- SIMILARITY: Contains 2 RNase III domains.

CC -1- SIMILARITY: Contains 1 PAZ domain.

CC -1- SIMILARITY: Contains 1 DRBM (double-stranded RNA-binding) domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC DR L14331; AAA28101.2; -.

CC DR WormPep; K12H4.8; CE25057.

CC DR InterPro; IPR001410; DEAD.

CC DR InterPro; IPR001159; DS\_RBD.

CC DR InterPro; IPR000534; DUF283.

CC DR InterPro; IPR001650; Helicase\_C.

CC DR InterPro; IPR003100; PAZ.

CC DR InterPro; IPR000999; RNase\_III.

CC DR Pfam; PF00270; DEAD; 1.

CC DR Pfam; PF00035; dsrm; 1.

CC DR Pfam; PF03368; DUF283; 1.

CC DR Pfam; PF00271; helicase\_C; 1.

CC DR Pfam; PF00636; Ribonuclease\_3; 2.

CC DR SMART; SMO0490; HELICC; 1.

CC DR SMART; SMO0358; DSRM; 1.

CC DR PROSITE; PS50137; DS\_RBD; 1.

CC DR PROSITE; PS50821; PAZ; 1.

CC DR PROSITE; PS00517; RNASE\_3-1; 1.

CC DR PROSITE; PS50142; RNASE\_3-2; 2.

CC KW Helicase; ATP-binding; RNA-binding; Hydrolase; Nuclease; Endonuclease; Repeat.

FT NP\_BIND 33 40 ATP (POTENTIAL).

FT DOMAIN 759 915 PAZ.

FT SITE 145 148 DECH BOX.

FT DOMAIN 1316 1524 RNASE III 1.

FT DOMAIN 1578 1740 RNASE III 2.

FT DOMAIN 1768 1831 DRBM.

SQ SEQUENCE 1845 AA; 210922 MW; 4A96EA2922FF1D9A CRC64;

Query Match 6.6%; Score 350; DB 1; Length 1845;

Best Local Similarity 23.5%; pred. No. 1.2e-09;

Matches 175; Conservative 118; Mismatches 249; Indels 202; Gaps 37;

Qy 309 RPKQMEVAQPALEGKNTTICLPTGSKTRVAVYIAKDH----LDKKKKASEPGKVV 3.61

Db 14 RDYQVELLDKATK-KNTIVVQLGTGSKTPFLAVLLKEYGVQLFAPLDQGGK----RAFF 67

Qy 362 LVNKVLLVEQLFRKEFQPFLLKKWYRVLIGLSDT----QLKLSFPEVVKSCDILLSTAQI 416

Db 68 VVERKVNVLVEQ---QAIHIEVHTSFKVGQVHGTGOTSSGLWDSKEQCDQFMKRHHVVVITAQC 124

Qy 417 LENSLNNLENGEDAGVQLSDFSLIILIDECHT-NKEAVYNNIMRHYLMQKLKNRJKKEN 475

Db 125 L----LDLIRHAYLKIEDMCVLIFDECHHALGSGQHYPYRSIMVDY----KLUKCD 170

Qy 476 KPVIPLPQILGLTASPVGGGATKQAKAEEHILKLCANLDAFTIKTVKENLDQLKNQIQEP 535

Db 171 KPV---PRVLLGTLASL-IAKVAPEKLMQLRKCLLESAMDS-VIETASD-LVLSLSKYGAKP 224

Qy 536 ----CKKFAIA--DATEDPFKEKLLIMTRIQTYCQMSPMSDFGTQ----- 576

Db 225 YEVVTLICKDFEIGCLGIPNPNEFTVIEFDETVAFVNNTTEFHPDLDLPRRPIKDSLKTTTR 284

Qy 577 ----PYEQWAIQMEKKAAKKGKGNRKERVCAELRKYNEALQIND-TIRMID-AYTHLE 627

Db 285 AVFRQLGWPAAW----RTAQWWEKELGKLIKSQVLPDKTRFLNMAKTSMI 331

Qy 628 TFYNEEKDKKFAVIEDDSDEGDDYECDGDEDDEDDLKKPLKLDETDRFTMFLFFENKML 687

Db 332 TI----KRLI-----EPEMKKKIKSIEALRPYVQPQRVIRLFIL 365

Qy 688 KRLAENPEYENERKLTKLRNTTIMEQYXTRTEESARGLIIFTKTRQSYAL----SQWITE 740

Db 366 ETF--NPEFQKRM-KLEKA-----EHLSSAIIFDQRYIAYSLLMMRHIKSW--- 410

Qy 741 NEKFAEVGVKAHHLIGAG---HSSEFKPMTQNEQEVISKFRTGKINILLIATTVAAEGL 796

Db 411 EPFKF--FVNPDYVVGASGRNLASSDSQGL-HKRQTEVLRFFHRNEINCLIASTSVEEGV 467

Qy 797 DIKECNIVRYGLVTNEIAMVQARGRARADESTYLVIAHSG-----SGVIE 842

Db 468 DVKQCNLVVIKEDRPLDMRSYVQSKGRARRAGSRYVITVEEKTAAAYSCKLPSDFITRLVP 527

Qy 843 HETVNDFRERMYKAIHCVNQNMKP--EEYAHKILELQIMQSIEMKMKTKRNIAKHYKNP 900

Db 528 HNQIIPIEENGWTK--YCAELLLPINSPIKHAI-----VLKNPMPNPKT----- 569

Qy 901 SLITFLCKNCVNSVLAACSGE-DIHVTEK-----FKE 937

Db 570 AQMVALEACRQLHLEGELDDNLPLKGRESIAKLEHIDEEPDEYAPGIAAKYVGSSKRKQ 629

Qy 938 LYIVRENKAL-----QKKCADY 954

Db 630 LYDKKIARALNESFVEADKECFIY 653

RESULT 7

ID Y1S2\_YEAST STANDARD; PRT; 993 AA.

AC P40562;

DT 01-FEB-1995 (Rel. 31, Created)

DT 01-FEB-1995 (Rel. 31, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Putative ATP-dependent RNA helicase Y1R002C.

GN Y1R002C OR Y1B2C.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycetaceae; Saccharomyces.

OC Saccharomyces cerevisiae (Baker's yeast).

OX NCBI\_TAXID=4932;  
 RN [1] SEQUENCE FROM N.A.  
 RP STRAINE=S288C;  
 RC MEDLINE=95282515; PubMed=7762303;  
 RX RA Voss H., Tamames J., Teodoru C., Valencia A., Sensen C., Wiemann S.,  
 RA Schwager C., Zimmermann J., Sander C., Ansorge W.;  
 RT "Nucleotide sequence and analysis of the centromeric region of yeast  
 RT chromosome IX.";  
 RL Yeast 11:61-78 (1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAINE=S288C / AB972;  
 RX MEDLINE=97313266; PubMed=9169870;  
 RA Churcher C.M., Bowman S., Badcock K., Bankier A., Brown D.,  
 RA Chillingworth T., Connor R., Devlin K., Gentles S., Hamlin N.,  
 RA Harris D.E., Horssell T., Hunt S., Jages K., Jones M., Lye G.,  
 RA Moule S., Odell C., Pearson D., Rajandream M.A., Rice P., Rowley N.,  
 RA Skelton J., Smith V., Walsh S., Whitehead S., Barrell B.G.;  
 RT "The nucleotide sequence of *Saccharomyces cerevisiae* chromosome IX.";  
 RL Nature 387:84-87 (1997).  
 CC - SIMILARITY: Belongs to the DEAD box helicase family. DEAH  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@isb-sib.ch).  
 DR EMBL; X79743; -; NOT\_ANNOTATED\_CDS.  
 DR EMBL; Z38062; CAN86204.1; -.  
 DR PIR; S48436; S48436.  
 DR GermOnline; 139714; -.  
 DR SGD; S0001441; MPFI.  
 DR GO; GO:0005634; C:nucleus; IDA.  
 DR GO; GO:0003724; RNA helicase activity; IMP.  
 DR GO; GO:0006281; P:DNA repair; IMP.  
 DR InterPro; IPR001410; DEAD.  
 DR InterPro; IPR002464; DEAD\_box.  
 DR InterPro; IPR001650; Helicase\_C.  
 DR Pfam; PF00270; DEAD; 1.  
 DR Pfam; PF00271; helicase\_C; 1.  
 DR SMART; SM00487; DEXDC; 1.  
 DR SMART; SM00490; DEAH\_ATP\_HELICASE; FALSE NEG.  
 DR PROSITE; PS00690; DEAH\_ATP\_BINDING; RNA-binding; Helicase.  
 KW Hypothetical protein; ATP-Binding; RNA-binding; Helicase.  
 FT NP BIND 107 114 ATP (POTENTIAL).  
 FT SITE 209 212 DEAH BOX.  
 SQ SEQUENCE 993 AA; 114057 MW; 474DDC99C543171F CRC64;  
 Query Match 5.7%; Score 303.5; DB 1; Length 993;  
 Best Local Similarity 20.3%; Pred. No. 1e-07;  
 Matches 143; Conservative 118; Mismatches 246; Indels 197; Gaps 25;  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX SEQUENCE FROM N.A.  
 RC STRAIN=168 / JH642;  
 RX MEDLINE=97124195; PubMed=8969508;  
 RA Mizuno M., Masuda S., Takemaru K.-I., Hosono S., Sato T., Takeuchi M.,  
 RA Kobayashi Y.;  
 RT "Systematic sequencing of the 283 kb 210 degrees-232 degrees region of  
 the *Bacillus subtilis* genome containing the *skin* element and many  
 RT sporulation genes.";  
 RT Microbiology 142:3103-3111 (1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=98044033; PubMed=9384377;  
 RA Kunst F., Ogabawara N., Moszer I., Albertini A.M., Alloni G.,  
 RA Berteiro M.G., Bessieres P., Bolotin A., Borchert S.,  
 RA Boriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,  
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,  
 RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,  
 RA Denizot F., Devine K.M., Dusterhoff A., Ehrlich S.D., Emmerson P.T.,  
 RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,  
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,  
 RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,  
 RA Guiseppi G., Guy B.J., Haga K., Hailech J., Harwood C.R., Henaut A.,  
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,  
 RA Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,  
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,

Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H., Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M., Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S., Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y., Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F., Sekiguchi J., Sekowska A., Serror S.J., Serror P., Shin B.S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K., Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A., Vian A., Wambutt R., Wedler E., Weitzenerger T., Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K., Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A., subtilis.”;	“The complete genome sequence of the Gram-positive bacterium <i>Bacillus subtilis</i> .”	Nature 390:249-256 (1997).	-!- SIMILARITY: Belongs to the SNF2/RAD54 helicase family.
<p>This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).</p>			
EMBL; D84432; BAA12545.1; -	PIR; G99116; CAB14389.1; -	SEQUENCE 557 AA; 64672 MW; 8693636CC219061A CRC64;	RP CHARACTERIZATION. MEDLINE=95196755; PubMed=7889940;
Subtilist; BG11699; yghH. InterPro; IPR001410; DEAD. InterPro; IPR000330; SNF2_N. InterPro; PF00271; helicase_C_1. PFam; PF00176; SNF2_N; T_1. SMART; SM00487; DEXDC_1. SMART; SM00490; HELIICC_1. Hypothetical protein; Helicase; ATP-binding; Complete proteome. NP BIND SITE 83 90 SITE 175 178 SITE DEAH BOX.	Query Match Best Local Similarity 21.5%; Pred. No. 0.00045; Matches 141; Conservative 98; Mismatches 198; Indels 218; Gaps 31;	RP CHARACTERIZATION. MEDLINE=95196755; PubMed=7889940; RX Thrower D.A., Jordan M.A., Schaar B.T., Yen T.J., Wilson L.; RA RA Yen T.J., Li G., Schaar B.T., Szilak I., Cleveland D.W.; RT RT "Mitotic HeLa cells contain a CENP-E-associated minus end-directed microtubule motor."; RT RT "CENP-E is a putative kinetochore motor that accumulates just before mitosis."; RT RT Nature 359:536-539 (1992). RN [2]	RP CHARACTERIZATION. MEDLINE=95196755; PubMed=7889940; RX Thrower D.A., Jordan M.A., Schaar B.T., Yen T.J., Wilson L.; RA RA Yen T.J., Li G., Schaar B.T., Szilak I., Cleveland D.W.; RT RT "Mitotic HeLa cells contain a CENP-E-associated minus end-directed microtubule motor."; RT RT Nature 359:536-539 (1992). RN [2]
SEQUENCE 557 AA; 64672 MW; 8693636CC219061A CRC64;	Query Match Best Local Similarity 21.5%; Pred. No. 0.00045; Matches 141; Conservative 98; Mismatches 198; Indels 218; Gaps 31;	RP CHARACTERIZATION. MEDLINE=95196755; PubMed=7889940; RX Thrower D.A., Jordan M.A., Schaar B.T., Yen T.J., Wilson L.; RA RA Yen T.J., Li G., Schaar B.T., Szilak I., Cleveland D.W.; RT RT "Mitotic HeLa cells contain a CENP-E-associated minus end-directed microtubule motor."; RT RT Nature 359:536-539 (1992). RN [2]	RP CHARACTERIZATION. MEDLINE=95196755; PubMed=7889940; RX Thrower D.A., Jordan M.A., Schaar B.T., Yen T.J., Wilson L.; RA RA Yen T.J., Li G., Schaar B.T., Szilak I., Cleveland D.W.; RT RT "Mitotic HeLa cells contain a CENP-E-associated minus end-directed microtubule motor."; RT RT Nature 359:536-539 (1992). RN [2]
304 PELQLRPYQMEYAAQPALEGKNN--IIIICLPTGSGKTRAVVYIAKDHLDKKKKASEPGKVIV 361 56 PSFTPLPHQIYEAQKVVEKMNGKAIIADEVGLGKTVEARGLIKEYMR---GLARKILLI 111	187 ---	RP FARNESYLATION. MEDLINE=98437347; PubMed=9763420; RX Chan G.K.T., Schaar B.T., Yen T.J.; RA RA Bishop W.R., Kirschmeier P.; RT RT "Characterization of the kinetochore binding domain of CENP-E reveals interactions with the kinetochore proteins CENP-F and hUBR1."; RL RL J. Cell Biol. 143:918-926 (1995). RN [3]	RP FARNESYLATION. MEDLINE=98437347; PubMed=9763420; RX Chan G.K.T., Schaar B.T., Yen T.J.; RA RA Bishop W.R., Kirschmeier P.; RT RT "Characterization of the kinetochore binding domain of CENP-E reveals interactions with the kinetochore proteins CENP-F and hUBR1."; RL RL J. Cell Biol. 143:918-926 (1995). RN [3]
362 LVNKVLLVEQLPRKEFQPFLL---KKWYRVIQGLSGDTQLKTSFPEVYKSCDIIISTAQ 415 112 LV-PASLVSQWVKELOQEKFLIPAVEQQKSY---WEQCDIVVSSID 153	416 ILENSLNLNGEDAGVQLS-DFSLIILIDECHHTNKEAVYNINMRYLMOKLKNRNLKKE 474 154 TAKRS-----PHREIVLSPYDVLVIDEAH-----KLKNNSKTK-- 186	RP FARNESYLATION. MEDLINE=98437347; PubMed=9763420; RX Chan G.K.T., Schaar B.T., Yen T.J.; RA RA Bishop W.R., Kirschmeier P.; RT RT "Characterization of the kinetochore binding domain of CENP-E reveals interactions with the kinetochore proteins CENP-F and hUBR1."; RL RL J. Cell Biol. 143:918-926 (1995). RN [3]	RP FARNESYLATION. MEDLINE=98437347; PubMed=9763420; RX Chan G.K.T., Schaar B.T., Yen T.J.; RA RA Bishop W.R., Kirschmeier P.; RT RT "Characterization of the kinetochore binding domain of CENP-E reveals interactions with the kinetochore proteins CENP-F and hUBR1."; RL RL J. Cell Biol. 143:918-926 (1995). RN [3]
475 NKPVTPLPQILQLTASPGVGGATKQAAKAEHHILKLCANLDAFTIKTYKXENLDQLKNQIQE 534 187 ---	535 PCKRFAIADATREDPFKKEKLLIMTRIQTYCQMSPMSPMSDFGTQPYEQWAQMEKKAQKGN 594 198 YC---LLLTAT---PIONRRIEEIFNLVSL---LKP-GHLGSQNHFQ---EEFAKKSS 242	CC -!- FUNCTION: MINUS-END DIRECTED MICROTUBULE MOTOR. PROBABLE KINETOCHEORE MOTOR. ACCUMULATES JUST BEFORE MITOSIS AT THE G2 PHASE OF THE CELL CYCLE. PROBABLY IMPORTANT FOR CHROMOSOME MOVEMENT AND/OR SPINDLE ELONGATION.	CC -!- SUBCELLULAR LOCATION: ASSOCIATES WITH KINETOCHEORES DURING CONGRESSION, RELOCATES TO THE SPINDLE MIDZONE AT ANAPHASE, AND IS QUANTITATIVELY DISCARDED AT THE END OF THE CELL DIVISION.
595 RKEVRVCAEHLRKYNEALQINDTIRMDA-----YTHLETFYNEEKDKKFAVIEDDSDEGG 649 243 LEAH---EHLKDLVNKVMDR---RRHDTGLNWKKQRHVETVPIQFSPEQALYDEIS--- 293	595 RKEVRVCAEHLRKYNEALQINDTIRMDA-----YTHLETFYNEEKDKKFAVIEDDSDEGG 649 243 LEAH---EHLKDLVNKVMDR---RRHDTGLNWKKQRHVETVPIQFSPEQALYDEIS--- 293	CC -!- SIMILARITY: Belongs to the kinesin-like protein family.	CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@isb-sib.ch).  
 CC

DR EMBL; Z15005; CAA78727.1; - .  
 DR HSSP; P17119; 3KAR.  
 DR Genew; HGNC:1856; CENPE.  
 DR GK; Q02224; - .  
 DR MIM; 117143; - .  
 DR GO; GO:0005699; C:kinetochore; TAS.  
 DR GO; GO:0005634; C:nucleus; TAS.  
 DR GO; GO:0008350; F:kinetochore motor activity; TAS.  
 DR GO; GO:0000067; P:DNA replication and chromosome cycle; TAS.  
 DR GO; GO:0007079; P:mitotic chromosome movement; TAS.  
 DR InterPro; IPR001752; kinetochore plate congression; TAS.  
 DR Pfam; PF00225; kinetochore\_motor.  
 DR PRINTS; PR00380; KINESINHEAVY.  
 DR SMART; SM00129; KISC; 1.  
 DR PROSITE; PS00411; KINESIN\_MOTOR\_DOMAIN1; 1.  
 DR PROSITE; PS50067; KINESIN\_MOTOR\_DOMAIN2; 1.  
 KW Motor protein; Cell division; ATP-binding; Coiled coil; Mitosis;  
 KW Cell cycle; Centromere; Lipoprotein; Prenylation.  
 FT DOMAIN 1 335  
 FT DOMAIN 336 2471  
 FT DOMAIN 2472 2663  
 FT NP\_BIND 86 93  
 FT LIPID 2660 2660  
 SQ SEQUENCE 2663 AA; 312087 MW; CEFC13880C8C8CB8 CRC64;

Query Match 4.1%; Score 217.5; DB 1; Length 2663;  
 Best Local Similarity 18.6%; Pred. No. 0.0072;  
 Matches 209; Conservative 185; Mismatches 421; Indels 307; Gaps 47;

Db 1097 -----QQEIVAQEKNAIKKEGELSRTCDRLAEVEKLEKSQLQEQKQQQLNNVQEE 1149  
 QY 9 ENFRLISCFRARVKMYIQQEPVLDYLTFPLPAEVKEQIORTVATSGNMQAVELLLSTLEK 68  
 Db 1040 EQQRKIFSLLIQEKNLNQMLLESVIAKEQLKTDLKENIEMTIE--NOEELRLLQGDELKK 1096  
 QY 69 GVWHLGWTREFV----EALRRTGSPSPLAARYMNPELTDLPSPSFENAHDYQLLNLQP 123  
 Db 124 TLVDKLLYRDV--LDKCMEEELLTIE--DRNRIAAEN-NGNESGVRELLK-RIYQKE-- 175  
 Db 1150 MSEMQKKINEIENLKNELKKNKEITLLEMETERLELAQKLNNENYEEVKSTTKVLKELQ 1209  
 QY 176 -----NWFSAFLNVLRQTG----NNELVQELTGSDCSESNAEIELN 212  
 Db 1210 KSFETERDHLRGYIREIEATGLQTKEELKIAHILKEBHQETIDELRRS-VSEKTAQILNT 1268  
 QY 213 SQVDGPQV----EQLQSLTTVQPNLEKEVWGMENNSSESSFADSSVVSSESDS 261  
 Db 1269 QDLEKSHTKLQEEIPVLHEEQELL----PNVKRVSSETQETMNLLELTQSTTKDS-TT 1322  
 QY 262 LAEGSVSCLDESIGHNSNMGSDEENVAARAPASPEPELQLRPYQMEVAQPALE 321  
 Db 1323 LARIEM---ERLRLNEKF----QESQEEIKSHTKASEPGKVIVLNKVLLEQ 1368  
 QY 322 GKNIIICLPTGSGKTRVAVYIAK-----DHLDKKKKASEPGKVIVLNKVLLEQ 372  
 Db 1369 -----KEHIRETLAKIQESQSKEQSLNMKEKDNETTKI-----VSEMEQ- 1408  
 QY 373 FRKEFQPPFLKKWYRIGLSDTQLKISFPE--VVKSCDIIISTAQILENSLJLNLENG-- 427  
 Db 1409 FKPKDSALLRIEIMGLS--KRLQESHDEMKSVAKEKDDLQRLOEVLOSESDDQLKENIK 1466  
 QY 428 EDAGVQLSDFSLIILIDECHHTNKEAVYNNIMRHYLMQKLKNNRHLKPKVPIPLPQILGL 487  
 Db 1467 EIVAKHLETEEELKVAHCLKEQETIN----ELRVNLSEKETE---- 1506

Qy	488	TASPGVGGATKOAKA-----EEHILKLCANILDAFTIKTVKENIDQLKNQIQEPCKKF	539
Db	1507	-----ISTIQQKLEAINDKLQNQKQIYEKEEQLNIKQISEQENVNELK-QFKEHRK--	1558
Qy	540	AIADATREDPFKEKLLIMTRIQTYCQ-MSPMSDFGTQPYEQWAIQMEKKAAKKGNR--	595
Db	1559	--AKDSALQSIESKOMLELTNRLQESQEEIQIMIKEKEEMKRVQEALQIERDQLKENTREI	1616
Qy	596	-----KERVCAEHLRKYNEALLQIN-----DTIRMID-AYT	624
Db	1617	VAKMKESQEKEYQFLKMTAVNETOQEMCEIEHLKEQFETQKLNLLENIEETIRLTLQILHE	1676
Qy	625	HLETTFYNBEXDKCKFAVIEDDSDEGGDDYCDGDEDDEDDLKLKPLKLD-----	670
Db	1677	NLEEMRSVTKER-----DDLRSVEETLKVVERDQLKENLRETTITRDLKEQELKIVHMH	1730
Qy	671	-----ETDTRFLMTLFFENNKMLKRLAENPEYENKL-----TKLRTNTIMEQYTRTE	716
Db	1731	KEHQETIDKLRLRGIVSEKTNETISNMQDLEHSNDALKAQDLKIQEELRIAHMLKEQQETI	1790
Qy	717	ESARGLIIFTKTRQSAYALSQWITEN---EKFAEVGVKAHHLIGAGHS-----SEPK	764
Db	1791	DKLRGIVSEKLDKLSNMQDLENSNAKLOKEKQELKANEHQLITKKDVNETOKVSEME	1850
Qy	765	PMTQ--NEQKEVISKFRTGKINLLIATTVAEEGLDIKECNIVIRGLVTNBIAQMARGR	822
Db	1851	QLKQIKDQSLTTSKLEIENLNL-AQELHENLEEMK-----SYMKERDN	1893
Qy	823	ARADESTYVILVAHSGSGV-----EH-ETVNDFEREKMYKAH	859
Db	1894	LRRVEETLKLERDQLKESLQETKARDLEIQQELKTARMLSKEHKETVDKLREKISEKTIQ	1953
Qy	860	CVQNMK-----PEEYAHKITLELQMQSIM-----EKRMKTKRNTIAKHYKNPNPLI	903
Db	1954	ISDIQKLDKSKDELQKKQELQKKELQDLMHKKINEMEQKKQFEPN-----	2009
Qy	904	TFLCKRNCSVLAQSGEDIHVIKMMHVNMTPEFKELYVRENK	945
Db	2010	YLCK-----CEMDNFQLTKKLHE-----SLEEIRIVAKER	2039
RESULT 10			
USO1 YEAST	AC	P25386;	
USO1 YEAST	DT	01-MAY-1992 (Rel. 22, Created)	
USO1 YEAST	DT	01-MAY-1992 (Rel. 22, Last sequence update)	
USO1 YEAST	DT	15-MAR-2004 (Rel. 43, Last annotation update)	
USO1 YEAST	DE	Intracellular protein transport protein USO1.	
USO1 YEAST	GN	USO1 OR INT1 OR YDL058W.	
USO1 YEAST	OS	Saccharomyces cerevisiae (Baker's Yeast).	
USO1 YEAST	OC	Eukaryota; Fungi; Ascomycota; Saccharomyces cerevisiae; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces.	
USO1 YEAST	OX	NCBI_TaxID=4932;	
USO1 YEAST	RN	SEQUENCE FROM N.A.	
USO1 YEAST	RC	STRAIN=X2180-1A;	
USO1 YEAST	RX	MEDLINE=91185402; PubMed=2010462;	
USO1 YEAST	RA	Nakajima H., Hirata A., Ogawa Y., Yonehara T., Yoda K., Yamasaki M.;	
USO1 YEAST	RA	"A cytoskeleton-related gene, uso1, is required for intracellular protein transport in Saccharomyces cerevisiae.", J. Cell Biol. 113:245-260 (1991).	
USO1 YEAST	RT	[2]	
USO1 YEAST	RT	SEQUENCE OF 782-1790 FROM N.A.	
USO1 YEAST	RA	Hostetter M.K., Herman D.J., Bendel C.M., McClellan M., Tao N., Kendrick K.E.; Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.	
USO1 YEAST	RL	[3]	
USO1 YEAST	RN	SEQUENCE OF 1-8 FROM N.A.	
USO1 YEAST	RA	Bai Y., Symington L.S.; Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.	
USO1 YEAST	RL		

CC -!- FUNCTION: Required for protein transport from the ER to the Golgi complex.  
 CC -!- LOCATION: CYTOPLASMIC. ASSOCIATED WITH INTRACELLULAR  
 CC MEMBRANES. PROBABLY PRESENT ON VESICLES OPERATIONAL BETWEEN THE  
 CC ER AND THE GOLGI COMPLEX.  
 CC -!- DOMAIN: The rodlike tail sequence is highly repetitive, composed  
 CC of an heptapeptide repeat pattern characteristic of alpha-helical  
 CC coiled coils. May form filamentous structures in the cell.  
 CC -!- SIMILARITY: BELONGS TO THE VDP/US01/YBL047C FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; X54378; CAA38253.1; -  
 DR EMBL; L03188; AAB00143.1; -  
 DR EMBL; U53668; AAB66659.1; -  
 DR GermOnline; 140300; -  
 DR SGD; S0002216; US01.  
 DR InterPro; IPR008938; ARM.  
 DR InterPro; IPR002017; Spectrin.  
 DR InterPro; IPR006955; US01\_p115\_C.  
 DR InterPro; IPR006953; US01\_p115\_head.  
 DR Pfam; PF04871; US01\_p115\_C; 1.  
 DR Pfam; PF04869; US01\_p115\_head; 1.  
 KW Transport; Protein transport; Golgi stack; Cytoskeleton; Coiled coil.  
 FT DOMAIN 1 724 GLOBULAR HEAD.  
 FT DOMAIN 725 1790 COILED COIL (POTENTIAL).  
 FT DOMAIN 465 487 CHARGED (HYPER-HYDROPHILIC).  
 FT DOMAIN 991 1790 DISPENSABLE FOR THE PROTEIN FUNCTION.  
 FT DOMAIN 1172 1786 ASP/GLU-RICH (ACIDIC).  
 FT CONFLICT 847 847 G > E (IN REF. 2).  
 FT CONFLICT 924 924 E > K (IN REF. 2).  
 FT CONFLICT 1253 1253 V > I (IN REF. 2).  
 FT CONFLICT 1319 1319 I > V (IN REF. 2).  
 FT CONFLICT 1461 1461 N > S (IN REF. 2).  
 FT CONFLICT 1581 1581 G > S (IN REF. 2).  
 FT CONFLICT 1600 1600 I > V (IN REF. 2).  
 FT CONFLICT 1661 1661 R > S (IN REF. 2).  
 FT CONFLICT 1772 1772 D > DEDDDE (IN REF. 2).  
 SQ SEQUENCE 1790 AA; 206424 MW; 6CE2B216E9FD4B18 CRC64;

Query Match 4.0%; Score 214.5; DB 1; Length 1790;

Best Local Similarity 19.3%; Pred. No. 0.0058;

Matches 180; Conservative 158; Mismatches 336; Indels 259; Gaps 37;

904 LIKDIAAKITEIKAINENLEEMKIQCNNLSSKEKEHI--SKELVEYKSRFQSHDNLVAKLT 961

89 -!- PLAARY--MNPPELTDLPSPSFENAHEYIQLNLQOPTLVKL-----LVRDV 134

962 EKLKSLANNYKDMQAENESLIKAVEEESKNESIQLSNL--QNKIDSMSQEKENFQIERGS 1019

135 LDKCMEEELLTTEDRNRRIAAEENNNGNEGVRELLKRVQENWFSAFLNVLQR-----T 188

1020 TEKNIEQLKKTTISDL-----EQTKEIITSKSDSSKDEYESQISLILKEKLETATT 1068

189 GNNELVQELTGSDCSENAEIJENLSQVDGPOQVEQLISTTVQPNLEKUVGMENNSSESS 248

1069 ANDENVNKL--SELTKTREELE-----AEILAAAYKNLKNEL----ETKLETS 1108

249 FADSSVVSESDTSLAEGSVSCLDESGLHNSMGSDSGTMGS - DSDEENVAARASPEPELO 307

1109 EKALKEVKENEEHLKEEKIQLKEKATETKQQLNSLRANLESKEHELDAA-----Q 1160

Db 308 LRPYQMEVAQPALEGKNIICLPL--TGSGKTRVAVYIAKDHLDKKKKASEPGKVTVLN 364

1161 LKKYEEQIANKERQYNEEISQLNDEITSTQOENESIKKKNDELEGEVKAMKS----- 1212

CC	365 KVLLVQLFRKEFOPFLKKWYRIGLSGDTQLKISFPEVVKSCDIIISTAQILENSLNL 424	QY	365 KVLLVQLFRKEFOPFLKKWYRIGLSGDTQLKISFPEVVKSCDIIISTAQILENSLNL 424
CC	1213 -----TSEEQSNLKK-----SEIDALNLQIQLKKNE - TNEASLLE----SI 1250	Db	1213 -----TSEEQSNLKK-----SEIDALNLQIQLKKNE - TNEASLLE----SI 1250
CC	425 ENGEDAGVQLSDFSLIITIDECHHTNKEAVYNNIMRHYLMQKLKNN----RLKKENK 476	QY	425 ENGEDAGVQLSDFSLIITIDECHHTNKEAVYNNIMRHYLMQKLKNN----RLKKENK 476
CC	1251 KSVSESETVKIKEKLQ----SELEDKLKASEDKNSKYLELKQSE 1299	Db	1251 KSVSESETVKIKEKLQ----SELEDKLKASEDKNSKYLELKQSE 1299
CC	477 PV-----IPLPQILGLTASPGVGGATKOAQAAEHTILKLCANLDAFTIKTVKEN 524	QY	477 PV-----IPLPQILGLTASPGVGGATKOAQAAEHTILKLCANLDAFTIKTVKEN 524
CC	1300 KIKEELDAKTTTELKIQLEKITNLSK-----AKEKSESELSSRLKKTSSER---KNAEEQ 1350	Db	1300 KIKEELDAKTTTELKIQLEKITNLSK-----AKEKSESELSSRLKKTSSER---KNAEEQ 1350
CC	525 LDQLKNOIQEPCCKFAIADATREDPFKEKLEIMTRIQTYCQMSPMDSFGTQPYEWAQI 584	QY	525 LDQLKNOIQEPCCKFAIADATREDPFKEKLEIMTRIQTYCQMSPMDSFGTQPYEWAQI 584
CC	1351 LEKLKNEIQLKQNAFE----KERKL----LNEGSSSTTQEYSEKINT 1389	Db	1351 LEKLKNEIQLKQNAFE----KERKL----LNEGSSSTTQEYSEKINT 1389
CC	585 MEKKAAKKGNRKBVRVCAEHLRKYNEALQINDTIRMIDAYTHLETFYNEKDKKPAVID- 643	QY	585 MEKKAAKKGNRKBVRVCAEHLRKYNEALQINDTIRMIDAYTHLETFYNEKDKKPAVID- 643
CC	1390 LEDDELRLQNLNENL-----KAREIDNTRSELEKVSLSNDLEEKQNTIKSLQDE 1439	Db	1390 LEDDELRLQNLNENL-----KAREIDNTRSELEKVSLSNDLEEKQNTIKSLQDE 1439
CC	644 -----DSDEGGDDEYC----DGDEDDEDDLKKPLKLDETDRFLMLTFFENRKLKJLAENP 694	QY	644 -----DSDEGGDDEYC----DGDEDDEDDLKKPLKLDETDRFLMLTFFENRKLKJLAENP 694
CC	1440 ILSYKDKITRNDEKLJLSIERDNKRDLESILKEQRLRAAQESKAKV----EEGLKKLEEE 1493	Db	1440 ILSYKDKITRNDEKLJLSIERDNKRDLESILKEQRLRAAQESKAKV----EEGLKKLEEE 1493
CC	695 -----EYENEKLTKLRNNTIMEQYTRTEESARGIIFTKTROSAYALSOIWITENEKFAEV 747	QY	695 -----EYENEKLTKLRNNTIMEQYTRTEESARGIIFTKTROSAYALSOIWITENEKFAEV 747
CC	1494 SKKAELEKSKEMMKKLESTIESNETELKSSMETI----RKSDEKLEQ----SKKSAAE 1544	Db	1494 SKKAELEKSKEMMKKLESTIESNETELKSSMETI----RKSDEKLEQ----SKKSAAE 1544
CC	748 GVKAHHLIGAGHSSEFKPMTQNEQKEVISKFRGKINLLIATTVAEEGLDIKECNIVIR- 806	QY	748 GVKAHHLIGAGHSSEFKPMTQNEQKEVISKFRGKINLLIATTVAEEGLDIKECNIVIR- 806
CC	1545 DIK-----NLQHEKSDLISR----INESEKDIEELKSKLRI 1576	Db	1545 DIK-----NLQHEKSDLISR----INESEKDIEELKSKLRI 1576
CC	807 -----YGLVTTNEIAMVQARGRARADESTVYLVAHSGSGVIEHETVNDFREKMMYKAH 859	QY	807 -----YGLVTTNEIAMVQARGRARADESTVYLVAHSGSGVIEHETVNDFREKMMYKAH 859
CC	1577 EAKSGSELETTVKQELNNAQEKIRINAENT----VLRSKLEIDIERELKDKQAE 1625	Db	1577 EAKSGSELETTVKQELNNAQEKIRINAENT----VLRSKLEIDIERELKDKQAE 1625
CC	860 CVQNMKPEEYAH---KILELQMQSIMEKKMKTK 889	QY	860 CVQNMKPEEYAH---KILELQMQSIMEKKMKTK 889
CC	1626 IKSNQEEKELLTSRLKELEQELDSTQRAQSE 1658	Db	1626 IKSNQEEKELLTSRLKELEQELDSTQRAQSE 1658
RESULT 11			
GCC2_HUMAN			
ID	GCC2_HUMAN STANDARD; PRT; 1583 AA.	AC	Q8IWJ2; Q15045; Q8THD3; Q9H2G8;
CC	TISSUE=Brain;	DT	10-OCT-2003 (Rel. 42, Created)
CC	CC	DT	10-OCT-2003 (Rel. 42, Last sequence update)
CC	CC	DT	15-MAR-2004 (Rel. 43, Last annotation update)
CC	CC	DE	GRIP and coiled-coil domain-containing protein 2 (Golgi coiled-coil protein GCC185) (CTCN tumor antigen sel-1) (CIL-associated antigen KW-11).
CC	CC	DE	GCC2 OR KIAA0336.
CC	CC	OS	Homo sapiens (Human).
CC	CC	OC	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
CC	CC	OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CC	CC	NCBI_TaxID	9606; KW-11).
CC	CC	RN	SEQUENCE FROM N.A. (ISOFORM 1).
CC	CC	RC	TISSUE=Brain;
CC	CC	RX	MEDLINE=97349984; PubMed=9205841;
CC	CC	RA	Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N., Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
CC	CC	RA	"Prediction of the coding sequences of unidentified human genes. VII. The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro." DNA Res. 4:141-150(1997).
CC	CC	RN	[2]
CC	CC	RX	SEQUENCE OF 1-448 FROM N.A. (ISOFORM 2).
CC	CC	RA	MEDLINE=22188422; PubMed=12200376;
CC	CC	RA	Krackhardt A.M., Witzenz M., Harig S., Hodil F.S., Zaulis A.J., Chessa M., Barrett P., Gribben J.G.;
CC	CC	RA	"Identification of tumor-associated antigens in chronic lymphocytic

RT Leukemia by SEREX.";  
 RL Blood 100:2123-2131(2002).  
 RP SEQUENCE OF 1-349 FROM N.A. (ISOFORM 1).  
 RC TISSUE=Brain;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altenschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaney S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schncher A., Schein J.E., Jones S.J.M., Marra M.A.,  
 RA "Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences";  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [4]  
 RP SEQUENCE OF 60-840 FROM N.A., AND TISSUE SPECIFICITY.  
 TISSUE=Testis;  
 RX MEDLINE=21143360; PubMed=11149944;  
 RA Eichmuller S., Usener D., Dummer R., Stein A., Thiel D.,  
 RA Schadendorf D.;  
 RT "Serological detection of cutaneous T-cell lymphoma-associated  
 antigens";  
 RL Proc. Acad. Sci. U.S.A. 98:629-634 (2001).  
 RN [5]  
 RP SUBCELLULAR LOCATION.  
 RX MEDLINE=22450606; PubMed=12446665;  
 RA Luke M.R., Kjer-Nielsen L., Brown D.L., Stow J.L., Gleeson P.A.;  
 RT "GRIP domain-mediated targeting of two new coiled-coil proteins,  
 GCC88 and GCC185, to subcompartments of the trans-Golgi network.";  
 RL J. Biol. Chem. 278:4216-4226 (2003).  
 CC -!- FUNCTION: Probably involved in maintaining Golgi structure (By  
 similarity).  
 CC -!- SUBCELLULAR LOCATION: Cyttoplasmic; peripheral membrane protein  
 associated with the trans-Golgi network.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=1;  
 CC IsoId=QBIWJ2-1; Sequence=Displayed;  
 CC IsoId=QBIWJ2-2; Sequence=VSP\_007733;  
 CC -!- TISSUE SPECIFICITY: Ubiquitous.  
 CC -!- DOMAIN: Extended rod-like protein with coiled-coil domains.  
 CC -!- SIMILARITY: Belongs to the golgin family.  
 CC -!- SIMILARITY: Contains 1 GRIP domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@isb-sib.ch).  
 CC DR EMBL; AB002334; BAA20794.1; -.  
 DR EMBL; AF432211; AAL99918.1; ALT\_INIT.  
 DR EMBL; BC03774; AAH3774.1; -.  
 DR EMBL; AF273042; AAG34902.1; ALT\_INIT.  
 DR Genew; HGNC:23218; GCC2.  
 DR InterPro; IPR000237; GRIP\_domain.  
 DR InterPro; IPR003345; M\_repeat.  
 DR InterPro; IPR002017; Spectrin.  
 DR 537 NSELEQVNELGGLEETLKEKDQNDQKLEKLMQMKVLSDEKEVLSAEVKSLYEENNKL 596  
 QY 586 -EKKAACK-----GNRKERVCAEHLRKYNEALQINDTIRMDAYTHLETFYNEEKDKK 637  
 DB 597 SSEKKQLSRDLEVELSQQKEDVILKEHITOLEKKQL-----MVEEQDNLNKLLNEQVQK 651  
 QY 638 FAV-----DDEYCD----GDEDEDDLK 664  
 DB 652 LPVKTQLYGFLEMSEVSEEDSEEKDVNVNLQAVGESLAKINEERCNIAFQRDEKVLELE 711

665 KPLKLDETDRF----LMLTF--FENNKMLKRLAENPEYENK----LTKLRLNTIMEQ 711  
 712 KEIKCLOQEESVQCEELKSLLRLDYEQEKVLLR-KELEIQTSEKEALQSDLLEMKNA--NE 768  
 712 YTRTEESARGIIIFTKTROSAYALSQWITENEKFAEVGKAHHLIGACHSSEFKPMTQNEQ 771  
 769 KTRLENQN---LLIQVEEVSQTCSKSEIHNEK-EKCFIKEH-----ENLKPPLLEQKE 816  
 772 -----KEVISKFRTGKINLLIATTVAEEGLD--IKECNIVIRYGLVTTNEIAMVQ 818  
 817 LRDRRAELILLKDSLAKSPSVKNDPLSSVKREKECK--EKEEKINKIKLVA 873  
 819 ARGRARADESTYVLAHSGSGVIEHETVNDFREKMMYKAIHCVQNMK----PEEYAHKI 873  
 874 VKAKKELLISSRK----ETQTVKEELESIRSLRSEKQDLSASMRDLIQAESYKNLL 922  
 874 LELOMQSIMEKMKTKRNLIAKHYKNNPNSLITFLCKN----CSVLACSGEDIHVIERNHHV 929  
 923 LEYERQS---EQLDVEKERANNFEHRIEDLTQLRNSTLQ CETINSDNEDL--LARIETL 977  
 930 NMTPEFKELYIVRENKAQKKCADCYQINGEILCK 963  
 978 OSNAKLLEVQLEVQRA--KAMVDEKEAFLQK 1009

## RESULT 12

## BRR2 YEAST

ID BRR2 YEAST STANDARD; PRT; 2163 AA.  
 AC P32639;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Pre-mRNA splicing helicase BRR2 (EC 3.6.1.-) (Protein Snu246).  
 GN BRR2 OR RSS1 OR SNU246 OR YER172C OR SYGP-ORF66.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OC NCBI\_TAXID=4932;  
 RN [1] TAXID=4932;

SEQUENCE FROM N.A.  
 STRAIN=S288C / AB972;

RX MEDLINE=97313264; PubMed=9169868;  
 RA Dietrich P.S., Mulligan J.T., Hennessy K.M., Yelton M.A., Allen E.,  
 RA Araujo R., Aviles E., Berno A., Brennan T., Carpenter J., Chen E.,  
 RA Cherry J.M., Chung E., Duncan M., Guzman E., Hartzell G.,  
 RA Hinicke-Smith S., Hyman R.W., Kayser A., Komp C., Lashkari D., Lew H.,  
 RA Lin D., Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P.,  
 RA Oh C., Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T.,  
 RA Smith V., Taylor P., Wei Y., Botstein D., Davis R.W.;  
 RT "The nucleotide sequence of *Saccharomyces cerevisiae* chromosome V.";  
 RL Nature 387:78-81(1997).  
 RN [2]  
 RP CHARACTERIZATION.  
 RC STRAIN=DBY473;  
 RX MEDLINE=96304576; PubMed=8722763;  
 RA Noble S.M., Guthrie C.;

RT "Identification of novel genes required for yeast pre-mRNA splicing by means of cold-sensitive mutations.";  
 Genetics 143:67-80(1996).  
 [3]  
 RN CHARACTERIZATION.

RX MEDLINE=96324408; PubMed=8670905;

RA Lauber J., Fabrizio P., Teigelkamp S., Lane W.S., Hartmann E.,  
 RA Luehrmann R.;

RT "The HeLa 200 kDa U5 snRNP-specific protein and its homologue in *Saccharomyces cerevisiae* are members of the DEXH-box protein family of putative RNA helicases.";  
 EMBBO J. 15:4001-4015(1996).

CC - FUNCTION: REQUIRED FOR PRE-mRNA SPlicing.  
 CC - SUBCELLULAR LOCATION: Nuclear (Potential).  
 CC - DOMAIN: Composed of two similar domains.

CC - SIMILARITY: Belongs to the helicase family. SKI2 subfamily.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).  
 EMBL; UI8922; AAB64699.1; --.  
 DR PIR; S50675; S50675.  
 DR GermOnline; 139249; --.  
 SGD; S0000974; BRR2.  
 GO; GO:0005682; C:snRNP U4; IDA.  
 GO; GO:0005688; C:snRNP U5; IDA.  
 InterPro; IPR003593; AAA\_ATPase.  
 InterPro; IPR001410; DEAD.  
 InterPro; IPR001650; Helicase\_C.  
 Pfam; PF00270; DEAD; 2.  
 Pfam; PF02271; helicase\_C; 1.  
 DR SMART; SM00382; AAA; 2.  
 DR SMART; SM00487; DEXDC; 2.  
 DR SMART; SM00490; HELICC; 1.  
 DR SMART; SM00611; SEC63; 2.  
 KW Hydrolase; Helicase; mRNA processing; mRNA splicing; Spliceosome; Nuclear protein; ATP-binding; Repeat.  
 FT DOMAIN 521 923 I.  
 FT DOMAIN 1370 1757 II.  
 FT NP\_BIND 521 528 ATP (POTENTIAL).  
 FT NP\_BIND 1370 1377 ATP (POTENTIAL).  
 FT SITE 634 637 DEIH BOX.  
 FT SITE 1474 1477 DDAH BOX.  
 SQ SEQUENCE 2163 AA; 246183 MN; DFAF7E3B7168D944 CRC64;  
 Query Match 3.9%; Score 204.5; DB 1; Length 2163;  
 Best Local Similarity 19.1%; Pred. No. 0.024;  
 Matches 162; Conservative 136; Mismatches 247; Indels 305; Gaps 40;

QY 127 DKLLYRDVLDKMEE----ELLTIED----RNR----IAAAE 156  
 DB 305 DTSVQDLSKEILNDIETLJEHNPVALEQKLVDLJKFENISLAERNSTIFWGIRLAK 364  
 QY 157 RNGNEGVRELKRVQKENWFAFLNVLRQTNNELVOELTGSDCSESNAEIENLSQVD 216  
 DB 365 STENI--IPNLTEKMYAK----GNDLVEQYKFRETTTHSKRELD--SGDD 406  
 QY 217 GPQVE---QLLSTTVQDNLEKEVGMENNSSESSFAESSFADSS--VVSSESDTSLAEGSVSCLD 271  
 DB 407 QPQSSEAKRTKFSNPALPPVIDE----KIKFDESSSKLMVTVKSLPEGSFK-- 454  
 QY 272 ESLGHNSNMGSDSGTMGSDSDEENVARASPEPELQLR----309  
 DB 455 ----RVKPQYDEIHIPASKPVIDYELKEITSLPLDWCOAFPSSETSTLN 500  
 QY 310 PYQMEVAQPALEG-KNLLICLPTGSGKTRAVY----IAKDHLDKKKKASEPGKVIVLV 363  
 DB 501 PIQSKVFAFAFEQGDSNMLICAPTGSGKTNIAILTVKLASHYNPKTKLNLSAFKIVYI 560  
 QY 364 N--KVLLVEQLFRKEFQ--PFLKKWYRIGLSSGTQKLKISPEVVKSCLIIISTAQILE 418  
 DB 561 APLKALVQEQQ--REFQRLAFL--GIVKVAELTGDLSRLS--RKQIDETQVLVSTPEKWD 613  
 QY 419 NSLLNLENEDAGVQLSDSFLIIIDECH--HTNKEAVYNNIM-RHYLMQKLKNNRKKEN 475  
 DB 614 ITTRNSNN--LAIVEL--VRLLIDIEIHLHDDGPVLESIVARTFWASKYQOE----663  
 QY 476 KPVIPPLPQIQLGILTASPVGGAATKQAAEEHILKLCANLDAFTIKTVKENLDQLKNOIQEP 535  
 DB 664 ----YPRIGLGSAT----LPNEYDVGFR-LRPVKEGLFYFDSSFR-P 700  
 QY 536 CKKFAIAADATREDPFKEKLEIMTRIOTYQWAIQMEKKAAKGNR 595

Db	701	C---	PLS-----QQFCGIGIKERNLSLKKLKA	721	NP BIND	3.5	42	ATP (POTENTIAL).		
Qy	5.96	KERVCAEH-LRKYNEALQINDTIRMIDAYTHLETFYNEEKDKKFAVIEDDSDEGGDEYC	654	FT DOMAIN	197	415	COILED COIL (POTENTIAL).			
Db	722	MNDACYEKVLSENEGNOI-----IVFVHSRKET	750	FT DOMAIN	446	1003	COILED COIL (POTENTIAL).			
Qy	655	DGDEDDEDDLKKPLKLDETDRFLMTLFPENNKMRLKRLAENPEYENELTKLNRNTIMEQYTR	714	SEQUENCE	1163	AA;	135507 MW;	CE5F0BD2215D7A92 CRC64;		
Db	751	-----SRTATWLKKNKFAEEN-----ITHKLTK	772	Query Match	3.7%	Score 197.5;	DB 1;	Length 1163;		
Qy	715	TEESARGIIFTKTKTROSAYALSQWITENERKFAEVGKKAHHLIGAGHSSEFFKPMTQNEQKEV	774	Best Local Similarity	18.8%	Pred. No. 0.022;	Mismatches 399;	Indels 263;	Gaps 45;	
Db	773	NDAGSKQIQLKT---EAANVLDPSL---RKLIESGIGTHH---AG-----LTERSDRSLS	816	Matches 197;	Conservative 187;	Mismatches 399;	Indels 263;	Gaps 45;		
Qy	775	ISKFRIGKINLLIATTVAAEGLDIKECNCNIVRYGLV-----TNEIAMVQARG	821	Query	31	VLDYLTF-LPDAEVKEQIQRITVATSGNMQAVELLSTLEKGWVHLLGWTREFVEALRRTGSP	89			
Db	817	EDLFADGLLQLVLCVATTLANGVNLPATHTVIIKGTDVYSPEKGWSWEQLSPDVLOMLGRAG	876	Db	44	ILDSITSLSYGEVARSKSSNFMNTNCNSLAVNSFEQFISGKETIKRYLVEREF----RRDNK	99			
Qy	822	RARADESTYVLAHSGSGVIEHETVNDFREKOMMYKAIHCVNQNMKPEEYAHKILLEQMQSI	881	Qy	90	LAARYMNPELTDLPSPSFENAHDEYLQQLNLLQPTLVDFKLLVRDVLDKCMEEBELLIEDR	149			
Db	877	RPRYD--TF----GEGII---ITDQSNVQYYLVLNQLPTEQFVSKLVDNLNAEV	924	Db	100	GSVRSKSAKIVDITGDEVEVLEEG-----AKSVNEKCQEIIIGLSLDDF	142			
Qy	882	MEKKMKTKRN 891	91	Qy	150	MRIAAEN-----NGNEGSVRELLKRVQKENW----FSAFLNVLRQ-TGNNELV	194			
Db	925	VAGNIKCRND 934	925	Db	143	TRTVVLPQGKSEFPLKLEGKER--RNMLERLFLNQLEYGDELSFKLARKKIRKEREKENV	200			
Qy	926	RESULT 13	SBCC CLOAB	Qy	195	QELTG----SDCSESENAELENLSQVDGQPQVEQOLLSTVQPNLEKEVWGMENNSSESSF	249			
Db	927	SBCC CLOAB	STANDARD	PRT;	1163	AA.	Db	201	GELKGYENTINEDVULKERRELLKENNDFFNEASKEYLKAEEEYNEGKEVWGLQIEEIEEKNR	200
Qy	928	SBCC CLOAB	Q97FKL;	Qy	250	ADSSVWSESEDTSLAEGSVSCLDESIGHNSNMGSDSGTMGSDSDEENVAARASPEPELQLR	309			
Db	929	SBCC CLOAB	Q97FKL;	Db	261	EIDLKEKARLG-ESSSKVK	289			
Qy	930	SBCC CLOAB	Q97FKL;	Qy	310	PY-----QMEVAQPALEGKNIITICLPTGSGKTRVAVYIAKDHLDDKKKASEPGKVI	360			
Db	931	SBCC CLOAB	Q97FKL;	Db	290	PYIDNYENTLQKIDILREQILSRE-----NTMKAISLEKEDMEKKLISIADNK	339			
Qy	932	SBCC CLOAB	Q97FKL;	Qy	361	VL-----VNKKVLLVEQLFRKEFQPFPLK-KWYRVIQLSGDTQLKISFEEV-----VKSCDII	410			
Db	933	SBCC CLOAB	Q97FKL;	Db	340	410 ALPKFMIKHHIILDAIKERDILIDNTIKLEKKRQLQKIEKLSLEASNKELIKONIKD	410			
Qy	934	SBCC CLOAB	Q97FKL;	Qy	411	410 ISTAQILENSLLNLENGEDAGVQLSD--FSLIIDE-CIHTN-----	449			
Db	935	SBCC CLOAB	Q97FKL;	Db	400	400 TLKTONLESKIDNLKVPPEEYKNNKTNNEGIFLRLNRYDEKLEKHQNKLGLDCDKFQVDFEAKS	459			
Qy	936	SBCC CLOAB	Q97FKL;	Qy	450	459 -KEAVYNNIMRHYLMQKLKNNRKKENKPKVITPLQILGLLTASPGVGA-TKQAKAAEELHIL	507			
Db	937	SBCC CLOAB	Q97FKL;	Db	460	507 KEMLFNKLEEEERSKLDTYTKKLQDLINKD-FPKDDVL-LTFQEKLNDSRQKWA	517			
Qy	938	SBCC CLOAB	Q97FKL;	Qy	508	517 KLCANLDAF----TIKTVKENLDQIKNQIQ----EPCKKKFAIADATRE-----	547			
Db	939	SBCC CLOAB	Q97FKL;	Db	518	547 SLKASLRVVENSEQVLRTKKEEMTKLEDKISKVNKIESLETENMAHVREKLKSGEACP	577			
Qy	940	SBCC CLOAB	Q97FKL;	Qy	548	577 -DPFKEKLEIMTRIQTYCQMSMSDF-GTQPYE----QWAQMEKCAA	590			
Db	941	SBCC CLOAB	Q97FKL;	Db	578	590 VCGSVHHIKEGFKEYDLKALETLK----SELEGFEKKRKPKPENEETVMCEASIRKVEE	632			
Qy	942	SBCC CLOAB	Q97FKL;	Qy	591	632 KGN-----YNEEK----DKRFAVIEDDSDEGGDDEDDLKKPLKLDDETDRFLMLTFPEN	683			
Db	943	SBCC CLOAB	Q97FKL;	Db	633	683 YQKEKTVKEQCEKRIVDLKSELEE-AIKEPNEAVYTENKALELKIQD----FKFEM	744			
Qy	944	SBCC CLOAB	Q97FKL;	Qy	684	744 FAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKFRGKINLLIATTVAEEGLDIKECNI	803			
Db	945	SBCC CLOAB	Q97FKL;	Db	794	803 -ABEL-----KEKDKIINEKIELI-KNKGVLNDLYELKEKIEG-TIKK--I	834			
Qy	946	SBCC CLOAB	Q97FKL;	Qy	804	834 VIRGLVTTNEIAMQARGRARADESTYVLAHSGSGVIEHETVNDFRERKMYKAIIHCVQN	863			
Db	947	SBCC CLOAB	Q97FKL;	Db	835	863 EQQNLCDKMMNEIEDKYBRCSD-----IUKYHNSLSSIIKDKPKYD	922			

RESULT 14	
b	GRIP and coiled-coil domain-containing protein 2 (Golgi coiled coil protein GCC185). GCC2 OR KIAA0336.
b	Mus musculus (Mouse) Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Murinae; Mus. [1]
b	SEQUENCE OF 1-769 FROM N.A. TISSUE=Breast cancer; MEDLINE=22388257; PubMed=12477932;
b	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klaunser R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmitz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E., Schanerch A., Schein J.E., Jones S.J.M., Marra M.A.; "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences." Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
b	SEQUENCE OF 1-442 FROM N.A. STRAIN=C57BL/6J; TISSUE=Brain, and Embryo; MEDLINE=22354683; PubMed=12466851;
b	Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S., Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H., Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T., Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J., Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W., Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S., Dalla B., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S., Gaasterland T., Garibaldi M., Gissi C., Godzik A., Gough J., Grimmond S., Gustincic S., Hirokawa N., Jackson I.J., Jarvis E.D., Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L., Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A., Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H., Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G., Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S., Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M., Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K., Sultana R., Takeda Y., Taylor M.S., Teasdale R.D., Tomita M., Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C., Wilming L.G., Wynnshaw-Boris A., Yanagisawa M., Yang I., Yang L., Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N., Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K., Shiraki T., Waki J., Aizawa K., Arakawa T., Fukuda S.,

Db	312 SVALQ-----EDPSAEQTVCDKVRQLEDSILKELESQHSILKDEVTYNNL----	356	OS Homo sapiens (Human).
QY	286 TMGSDSDEENVAARASPEPE-----LQLRFYQMEVAQPALEGKNTIICLPTGS	333	OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Db	357 KLKLEMDAQHKDEFERDLEFFKINELLILLAKEEQGYVVEKLKYEREDLNLNQLCCAVEQ	416	OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
QY	334 GKTRVAVYIAKDHLDKKKKASEPGKVIVLYNKVLLVEQLFRKEFQPFLKKWYRIGLSGD	393	NCBI_TaxID=9606;
Db	417 HNKEIQ-RLQEH-----QKEVSE-----LSETFISGSEKEKLALMFETOGLKEQ	460	[1] RN SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
QY	394 TQ-LKISFPEVVKSCDIIISTAQIILENSLLNLENGEDAGVQLSDFSLIIDECHHTNKEA	452	RX MEDLINE=961215236; PubMed=8626529;
Db	461 CENLQHEKQEVVLNYESLREMMEIQLTEL-----GESAGKISQEF-----ETMKQQ	506	RA Erlich R., Gleeson P.A., Campbell P., Dietzsch E., Toh B.-H.,
QY	453 VYNNIMRHRYLMQKLN-----NRLKKENKPVIPLPQIILGLTASPGVGATKQA	500	RT "Molecular characterization of trans-Golgi p230: a human peripheral membrane protein encoded by a gene on chromosome 6p12-22 contains extensive coiled-coil alpha-helical domains and a granin motif.";
Db	507 QASDV-HELQQKLRSAFNEKDALLETVNRLQGENEKLLSQQELV-----PELESTIENL	559	RT J. Biol. Chem. 271:8328-8337(1996).
QY	501 KA-----EHLKLICANLDAF-----TIKTVKEMLDQL-----KNIQIQ	533	RN [2] RN SEQUENCE FROM N.A.
Db	560 QADNSMXYLASLGQKDTMLOQELEAKISSLAKEKDDFISKITSHEEMDDLHQKWEREQRLS	619	RC TISSUE=Placenta;
QY	534 EPCKKFAIADATREDPEKEKLLIMTRIQTYCQMSPMSD-----FGTQPYEWAQIOMER	587	RX MEDLINE=96125112; PubMed=8537393;
Db	620 VELREAAGQAAQHNSLQRQVSELTGKLDELVREKSONDQSITVQMKTMTEDEQEALESSKI	679	RA Fritzler M.J., Lung C.-C., Hamel J.C., Griffith K.J., Chan E.K.L.,
QY	588 KAAKGNRKRKERVCAEHLRKYNEALQINDTIRMIDA-----YTHLETFYNEBKKFAVTEDD	644	RT "Molecular characterization of golgin-245, a novel Golgi complex protein containing a granin signature.";
Db	680 KSLYEEENN-----RLHSEKAQLSRDLALQAQODFAHKEHVAEEFEKKLQLMV-----	726	RL Thesis (1994). Instituto municipal de investigacion medica, Spain.
QY	645 SDEGGDDEYCDGDEDDEDDLKPKLDETDR-FLMLTLEFFNNKMLKR-----LAENPEYENEKL	701	CC -!- FUNCTION: May play a role in vesicular transport from the trans-Golgi.
Db	727 -----EERDDLNKLJLNEQVQKSFVKTQLYELYKLQRLASTILEEN-----EEEDV	770	CC -!- SUBCELLULAR LOCATION: Cytoplasmic, peripheral membrane protein associated with the trans-Golgi network.
QY	702 TKLRNNTIMEQYTRTEESARGLILFTKTROSAYALSQWITENEKFAEVGVKAHILLIGAGHSS	761	CC -!- ALTERNATIVE PRODUCTS:
Db	771 VKLIIQAVGESLKVREEEHNLFV-----EYDARVLELENK-----IKCLOQEDSAVQCE	818	CC -!- Event=Alternative splicing; Named isoforms seem to exist;
QY	762 EFKPMTQNEQKEVISKFRGKINL-----LIATTVAEEG-----LDIKECNIVTRYGLVTNEIAM	816	CC -!- Comment=Additional isoforms seem to exist;
Db	819 ELRTLTVRDESEOE-----KILLRKELDAVTSAKEALQDILLEMK-----VNDER-----	850	CC -!- IsoId=Q13439-1; Sequence=VSP_004272, VSP_004273;
QY	817 VQARG-RARADESTYVLAHSGSG-----VIEHET-----VNDER-----	850	CC -!- IsoId=Q13439-3; Sequence=VSP_004274;
Db	864 LENQLTSTQVEELSQTLL-HSRNEVHDEKVLVIEHENLRLIKKORESELDQVRAELILK	921	CC -!- IsoId=Q13439-4; Sequence=VSP_004275;
QY	851 -----EKMYYKAITHCVQNMKPEEYAHKI	873	CC -!- DOMAIN: Extended rod-like protein with coiled-coil domains.
Db	922 DSLEKSPSPVKDQLSLVKEEKKIESLEKESKDKEKISKIKLVAVKAKKELDSNRKEAQ	981	CC -!- DISEASE: Antibodies against GOLGA4 are present in sera from patients with Sjogren's syndrome (SS) [MIM:270150]. Sera from patients with SS often contain antibodies that react with normal components of the Golgi complex.
QY	874 LEILOMQSIMEKK-----MKTKRNIAKHYKCNPNSLITFLCKNCNSVLAGSGEDIHVIEKMH	927	CC -!- DISEASE: Antibodies against GOLGA4 are found in sera from hepatitis B patients.
Db	982 LREEELESVRSEKDRLSASMKEFLQGAESYK-----SLLLEYDKQSEQIDVEKERAHNFER--	1036	CC -!- SIMILARITY: Belongs to the golgin family.
QY	928 HV-----NMTPPEFKELYTVRENKALQKKCADYQINGEI-----ICKCGQAWGTMVHK	975	CC -!- This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>
Db	1037 HIEDLTKQLRNSTCQYERL-TSDNEDLLARIELTQANAKLLEAQILEVQAKKG--VVEK	1092	CC -!- DR EMBL; U41740; AAC50434.1; -.
QY	976 GLDLPC-----KIRNFV-----VVFKNISTKKQKQKWE	1005	CC -!- DR EMBL; X82834; CAA58041.1; -.
Db	1093 ELDAEELQKEQKIKEHVSTVNEELQFQKEKQLQKTMQ	1134	CC -!- DR EMBL; U31906; AAC51791.1; -.
			DR EMBL; X76942; CAA54261.1; -.
			DR Genew; HGNC:4427; GOLGA4.
			DR MIM; 602509; -.
			DR MIM; 270150; -.
			DR GO; GO:0005802; C:Golgi trans face; TAS.
			DR GO; GO:0016192; P:vesicle-mediated transport; TAS.
			DR InterPro; IPR000237; GRIP_domain.
			DR Pfam; PF01465; GRIP_1.
			GOA4_HUMAN RESULT 15
			ID GOA4_HUMAN STANDARD; PRT; 2230 AA.
			AC Q13439; Q13270; Q14654; Q14436;
			DT 16-OCT-2001 (Rel. 40, Created)
			DT 10-OCT-2003 (Rel. 42, Last annotation update)
			DE Golgi autoantigen, golgin subfamily A member 4 (Trans-Golgi p230) (72.1 protein).
			DE kDa golgin (Golgin-245) (72.1 protein).
			GN GOLGA4.

DR PROSITE; PS50913; GRIP; 1.  
 KW Antigen; Golgi stack; Coiled coil; Alternative splicing.  
 FT DOMAIN 133 2185 COILED COIL (POTENTIAL).  
 FT DOMAIN 252 2095 GLU-RICH.  
 FT DOMAIN 2168 2215 GRIP.  
 FT VARSPLIC 2154 2185 TPYKGGNLYHTDVSILFGEPEFEYLRKVLFY -> HLTKV  
 FT ACTTIRMSHSLLENLPNLSICEKCFLSI (in isoform 2).  
 FT VARSPLIC 2186 2230 /FTId=VSP 004272.  
 FT VARSPLIC 2103 2109 Missing (in isoform 2).  
 FT VARSPLIC 2222 2230 /FTId=VSP 004273.  
 FT VARSPLIC 2222 2230 Missing (in isoform 3).  
 FT VARSPLIC 2222 2230 /FTId=VSP 004274.  
 FT VARSPLIC 2222 2230 FTSPRSGF -> SWLRSSS (in isoform 4).  
 FT VARSPLIC 2222 2230 /FTId=VSP 004275.  
 FT CONFLICT 188 188 R -> K (IN REF. 3).  
 FT CONFLICT 220 220 Y -> H (IN REF. 3).  
 FT CONFLICT 276 276 T -> A (IN REF. 3).  
 FT CONFLICT 584 584 K -> E (IN REF. 3).  
 FT CONFLICT 628 628 T -> A (IN REF. 3).  
 FT CONFLICT 630 630 K -> E (IN REF. 3).  
 FT CONFLICT 682 682 K -> N (IN REF. 3).  
 SQ SEQUENCE 2230 AA; 261139 MW; 3BB733DB1EA86134 CRC64;  
 Query Match 3.7%; Score 197; DB 1; Length 2230;  
 Best Local Similarity 17.9%; Pred. No. 0.058;  
 Matches 185; Conservative 212; Mismatches 366; Indels 270; Gaps 46;  
 Qy 15 ISCFRARVRMYYIQVEPVYLDYLTFLPAAEVKEQIORTVATSGNMQAVELLSTLEKGWVHIG 74  
 Db 384 IAQLRSRIKQ----MTTQGEELRQEKEKSERAA--FEELEKALSTAQQ----425  
 Qy 75 WTRFVEALRRRTGSPLAARTMNPPELTDLPSPSFENAHDYELQQLNLLQPTLVDKLVLVRDV 134  
 Db 426 -TEE----ARRKLKAEM-DEQIKTTEKTSEE--ERISLQQELSRSVQEVVVDV 469  
 Qy 135 LDRCMEEELLTIEDRNRRIAAENNGNESGURELLKRIIVQKENWFAFLNVLRTQTNNELV 194  
 Db 470 MKKSSEEQIAKLOKLHEKELARKE----QELTKKLQLTREREQEQMKVALEKQSEYL 523  
 Qy 195 QELTGSDCSESNAEILNLSQVQDGPQVEEQLLSTTVQPNLKEVWGMENNSSESSSFADSSV 254  
 Db 524 K----ISQEKEQQESLA-LEEELQKKALLT-----ESENKLRLDQQ 560  
 Qy 255 VSESDTSLAEGSVSCLDESIGHNSNMGSDSGT-MGSDSIDEENVAARASPE--PELQLRP 310  
 Db 561 EAETYRTRILELESSLEKSLQENKNQSKDLAVHLEAKKHNEITVMVKHKTTELSLK 620  
 Qy 311 YOMEVAQPALEGKNNIICLPTGSKTRVAVIAKDHLKKKASEPQKVIVLNVKLL--  
 Db 621 HQ----QDAL----WTEKLQVLKQYQOTEMEKLREKCEQEKETLKDKOETIIFQ 665  
 Qy 369 -VEQLFRKEFQPFLKQWYRVIQGSDTQLKISPEVVKSCDIIISTAQILENSLLNLEN 426  
 Db 666 AHIEEMNEKETLEKLDVQTELESISSE----LSEVLKA----ESENKLRLDQQ 560  
 Qy 427 GEDAGVQLSDFSLLIIDECHHTNKEAVYNNIIMRHLYMQLKLNRLKENKPVIFLPQILG 486  
 Db 713 QTDKMKQELAKM--DEQAKNHQQQVDLSIKEH----EVSIQRTKEALKDQINQLELL- 764  
 Qy 487 LTASPGVGGATKQAKAAEEHILKLCANLDAFTIKTVKENLDQLKQIQEPCKKFAI----A 542  
 Db 765 ----LKBFRDKHLKEHQAH----VENLEADIKRSEGEQELQASAKLDFQSYQ 807  
 Qy 543 DATRED--PFKEKLLIMTRIQTQYQMSMPSDFGTQPYEWAQMEKKAAKGKRNKERV 600  
 Db 808 SATHEQTKAYEEQLAQLQQK----LIDOLETE----RILITKQVAEVEAKKKDVC 853  
 Qy 601 AE--HLRKYNEALQINDTIRMDAYTHLETFVNEDKDKFAVIEDDSDEGGDDEYCDGD 657  
 Db 854 TELDAH----KIQVQDLMQOLEK----QNSEMEOQVKVSLTQVYESKLE----DGN 896  
 Qy 658 EDEDLKKPLKLDDETDRFLMLTLEFFENNKKMLKRLAE----NPEYE---- 697

GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2004, 19:06:53 ; Search time 49 Seconds

(without alignments)  
 6600.126 Million cell updates/sec

Title: US-09-515-363C-2  
 Perfect score: 5311  
 Sequence: 1 MSNGYSTDENFRLISCFRA.....LPITFPNLDYSECCLFSDDED 1025

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL 25:\*

1: sp\_archea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_mhc:\*

8: sp\_organelle:\*

9: sp\_phage:\*

10: sp\_plant:\*

11: sp\_rat:\*

12: sp\_virus:\*

13: sp\_vertebrate:\*

14: sp\_unclassified:\*

15: sp\_rvirus:\*

16: sp\_bacteriaph:\*

17: sp\_archeap:\*

Database :

17: 687 12.9 1037 5 O44165 caenorhabdi  
 18: 678 12.8 143 11 Q8ve79 mus musculu  
 19: 555.5 10.5 472 11 Q8C7T2  
 20: 555 10.5 1164 5 Q93413 caenorhabdi  
 21: 516 9.7 398 4 Q9NT04 homo sapien  
 22: 516 9.7 620 5 Q17545 caenorhabdi  
 23: 453 8.5 752 17 Q9V1Z5 pyrococcus  
 24: 441.5 8.3 650 17 Q59524 pyrococcus  
 25: 431.5 8.1 1883 10 Q8LMR2 oryza sativ  
 26: 417 7.9 741 17 Q28814 archaeoglob  
 27: 395.5 7.4 764 17 Q8TZH8  
 28: 391 7.4 410 11 Q8C5I3 mus musculu  
 29: 389 7.3 821 17 Q8TUH1 methanosa  
 30: 384 7.2 1604 10 Q7XQ14 oryza sativ  
 31: 368 6.9 864 17 Q8PX35  
 32: 365.5 6.9 837 5 Q86L44 dictyosteli  
 33: 350 6.6 513 17 Q978A0 thermoplasm  
 34: 343 6.5 182 11 Q9D1X4  
 35: 343 6.5 2042 10 Q9M9P8  
 36: 342 6.4 738 17 Q27466  
 37: 332.5 6.3 1383 5 Q9VDA0  
 38: 332 6.3 508 17 Q9HT46  
 39: 313.5 5.9 784 17 Q9HMW5  
 40: 309.5 5.8 1883 5 Q95ZG7  
 41: 302.5 5.7 741 17 Q8TUS6  
 42: 302.5 5.7 1789 5 QBT145  
 43: 295.5 5.6 783 3 Q9HE09  
 44: 295 5.6 1458 10 Q8W367  
 45: 295 5.6 1458 10 Q7XD96

#### ALIGNMENTS

#### RESULT 1

Q9H3G6 ID: Q9H3G6 PRELIMINARY; PRT; 1025 AA.  
 AC: Q9H3G6;  
 DT: 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT: 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT: 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
 DE: Melanoma differentiation associated protein-5.  
 GN: MDA5.  
 OS: Homo sapiens (Human).  
 OC: Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC: Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 NCBI\_TAXID=9606;  
 [1]  
 RN: RP  
 RP: SEQUENCE FROM N.A.  
 RC: TISSUB=Melanoma;  
 RA: Kang D.-C., Fisher P.B.;  
 RT: "Melanoma differentiation associated gene-5 (mda-5), an interferon  
 inducible gene of limited homology to RNA helicase;"  
 RL: Submitted (SEP-1998) to the EMBL/GenBank/DDBJ databases.  
 DR: EMBL; AF095844; AAG34368.1;  
 DR: GO; GO:0005524; F:ATP binding; IBA.  
 DR: GO; GO:0008026; F:ATP dependent helicase activity; IBA.  
 DR: GO; GO:0016787; F:hydrolase activity; IBA.  
 DR: GO; GO:0003676; F:nucleic acid binding; IBA.  
 DR: InterPro; IPR001410;  
 DR: InterPro; IPR001650; Helicase\_C.  
 DR: Pfam; PF00270; DEAD; 1.  
 DR: Pfam; PF00271; helicase\_C; 1.  
 DR: SMART; SM004487; DEXDC; 1.  
 DR: SMART; SM00490; HELIC\_C; 1.  
 KW: ATP-binding; Helicase; Hydrolase.  
 SQ: SEQUENCE 1025 AA; 116687 MW; 48BE75491DB63741 CRC64;

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	5311	100.0	1025	4	Q9H3G6		Q9h3g6 homo sapien
2	5285	99.5	1025	4	Q9BYX4		Q9byx4 homo sapien
3	4199	79.1	1025	11	Q8R5F7		Q8r5f7 mus musculu
4	4194	79.0	1025	11	Q8K5C7		Q8k5c7 mus musculu
5	2673	50.3	693	11	Q8BZ01		Q8bz01 mus musculu
6	2633	49.6	683	11	Q8BYC9		Q8byc9 mus musculu
7	2456	46.2	468	4	Q96MX8		Q96mx8 homo sapien
8	2271.5	42.8	514	11	Q99KS4		Q99ks4 mus musculu
9	2059.5	38.8	467	11	Q8R144		Q8r144 mus musculu
10	1356	25.5	678	11	Q99J87		Q99j87 mus musculu
11	1352	25.5	678	4	Q96C10		Q96c10 homo sapien
12	1344	25.3	678	4	Q9HAM6		Q9ham6 homo sapien
13	1122	21.1	244	11	Q9D2Z5		Q9d2z5 mus musculu
14	1086.5	20.5	925	4	Q95786		Q95786 homo sapien
15	1071	20.2	221	4	Q86X56		Q86x56 homo sapien
16	1018	19.2	940	6	Q9GLV6		Q9glv6 sus scrofa

Query Match 100.0%; Score 5311; DB 4; Length 1025;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-288;  
 Matches 1025; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MSNGYSTDENFRLISCFRAVVKMYIQVEPVLDYLTLFLPAEVKEQIORTVATSGNMQAVE 60

Db	1	MSNGYSTDENFRYLISCFARVKMVIQVEPVLDYLTLPAEVKEIQTORTVATSGNMQAVE	60	Q9BYX4	PRELIMINARY;	PRT;	1025 AA.
QY	61	LLLSTLEKGWHLGWTREFVEALRRRTGSPLAARYMNPFLTDLPSPSFENAHDEYQLLLNL	120	Q9BYX4;			
Db	61	LLLSTLEKGWHLGWTREFVEALRRRTGSPLAARYMNPFLTDLPSPSFENAHDEYQLLLNL	120	AC			
QY	121	LQPTLVDKLLVVDLTDKCMEEBLLTIEDRNRRIAAENNGNGVRELLKRVYQKENWFA	180	DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)		
Db	121	LQPTLVDKLLVVDLTDKCMEEBLLTIEDRNRRIAAENNGNGVRELLKRVYQKENWFA	180	DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)		
QY	181	FUNVLRQTGNNELVQELTGSDCSESNAELENLSQVDGPPQVEEQLLSTTVQPNLEKEVWGM	240	DE	RNA helicase-DEAD box protein RH116.		
Db	181	FUNVLRQTGNNELVQELTGSDCSESNAELENLSQVDGPPQVEEQLLSTTVQPNLEKEVWGM	240	OS	Homo sapiens (Human).		
QY	241	ENNSSESSFAADSSVVSSESDTSLAEGSVSCLDESGLHNSNMGSDETMGSDSGTMGSDEENVAARA	300	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
Db	241	ENNSSESSFAADSSVVSSESDTSLAEGSVSCLDESGLHNSNMGSDETMGSDSGTMGSDEENVAARA	300	OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
QY	301	SPEPELQLRQPYQMEVAQPALEGKNNIIICLPTGSGKTRAVYIAKDHLDKKKASEPGKVI	360	OX	NCBI_TaxID:9606;		
Db	301	SPEPELQLRQPYQMEVAQPALEGKNNIIICLPTGSGKTRAVYIAKDHLDKKKASEPGKVI	360	RN	SEQUENCE FROM N.A.		
QY	361	VLYNKVLLVEQLFRKEFQPFKLKWYRVLIGLSGDTQLKISFPEVVKSCDIIISTAQILENS	420	RP	Cocude C., Kolesnichenko V., Billaut-Mulot O., Truong M.-J.,		
Db	361	VLYNKVLLVEQLFRKEFQPFKLKWYRVLIGLSGDTQLKISFPEVVKSCDIIISTAQILENS	420	RA	Capron A.; Bahr G.M.;		
QY	421	LLNLENGEDAGVQLSDFSLIIIDECHHTNKEAVNNIMRHYLMOQLKNRLLKVNKPV1P	480	RA	"Identification of a new RNA helicase (RH116) regulated by the		
Db	421	LLNLENGEDAGVQLSDFSLIIIDECHHTNKEAVNNIMRHYLMOQLKNRLLKVNKPV1P	480	RT	immunomodulator Murabutide."		
QY	481	LFDQILGLTASPVGGGATKQAKAEEHILKLCANLDAFTIKTVKENIDQLKNQIQEPCKKFA	540	RL	Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.		
Db	481	LFDQILGLTASPVGGGATKQAKAEEHILKLCANLDAFTIKTVKENIDQLKNQIQEPCKKFA	540	DR	EMBL; AY017378; AAG54076.1;		
QY	541	IADATREDPFKEKLLIMTRIQTQYQCMSPMSDFGTQTOPYQWAIQMEKKAQKGNRKERVY	600	DR	GO: 0005524; F:ATP binding; IEA.		
Db	541	IADATREDPFKEKLLIMTRIQTQYQCMSPMSDFGTQTOPYQWAIQMEKKAQKGNRKERVY	600	DR	GO: 0008026; F:ATP dependent helicase activity; IEA.		
QY	601	AEHLRKYNEALQINDTIRMTIDAYTHTLETFYNEEKDKKFAVIEDDSDEGGDDEYCDGDEDE	660	DR	GO: 0016787; F:hydrolase activity; IEA.		
Db	601	AEHLRKYNEALQINDTIRMTIDAYTHTLETFYNEEKDKKFAVIEDDSDEGGDDEYCDGDEDE	660	DR	InterPro; IPR001410; DEAD.		
QY	661	DDLKCKPLKLDETDRFLMTLFFENNNKMLKRLAENPEYENELKTLKLNTTIMEQYTRTEESAR	720	DR	InterPro; IPR001650; Helicase_C.		
Db	661	DDLKCKPLKLDETDRFLMTLFFENNNKMLKRLAENPEYENELKTLKLNTTIMEQYTRTEESAR	720	DR	Pfam; PF00270; DEAD; 1.		
QY	721	GIIFTKTRQSYALSQWITENEKFAEVGVKAHLIGAHSSEFKPMTQNEQEVISKFR	780	DR	SMART; SM00487; DEXDC; 1.		
Db	721	GIIFTKTRQSYALSQWITENEKFAEVGVKAHLIGAHSSEFKPMTQNEQEVISKFR	780	DR	SMART; SM00490; HELICC; 1.		
QY	781	GRINLLIATTVAEEGLDIKECNIVGLVTINELAMVQAGRRADESTYVLVAHSGSGV	840	KW	ATP-binding; Helicase; Hydrolase.		
Db	781	GRINLLIATTVAEEGLDIKECNIVGLVTINELAMVQAGRRADESTYVLVAHSGSGV	840	SQ	SEQUENCE 1025 AA; 116671 MW; 6B4F3D14E88A7D6D CRC64;		
QY	841	IEHETVNDFREKMMYKA1HCVQNMKPEEYAHKILELQMOSIMEKKMKT1KRN1AKHYKNNP	900	Query	Query Match 99.5%; Score 5285; DB 4; Length 1025;		
Db	841	IEHETVNDFREKMMYKA1HCVQNMKPEEYAHKILELQMOSIMEKKMKT1KRN1AKHYKNNP	900	Best Local Similarity 99.5%; Pred. No. 7.7e-287; Matches 1020; Conservative 2; Mismatches 3; Indels 0; Gaps 0;			
QY	901	SLITFLCKNCSVLAQSGEDIHVIKEKCNIVGLVTINELAMVQAGRRADESTYVLVAHSGSGV	960	Db	1 MSNGYSTDENFRYLISCFARVKMVIQVEPVLDYLTLPAEVKEIQTORTVATSGNMQAVE 60		
Db	901	SLITFLCKNCSVLAQSGEDIHVIKEKCNIVGLVTINELAMVQAGRRADESTYVLVAHSGSGV	960	Db	1 MSNGYSTDENFRYLISCFARVKMVIQVEPVLDYLTLPAEVKEIQTORTVATSGNMQAVE 60		
QY	961	ICKCQGQAWGTMVHKGLDLPCLKIRNFVUVFKONSTKQYKKWVELPITFPNLDYSECC1	1020	QY	1 MSNGYSTDENFRYLISCFARVKMVIQVEPVLDYLTLPAEVKEIQTORTVATSGNMQAVE 60		
Db	961	ICKCQGQAWGTMVHKGLDLPCLKIRNFVUVFKONSTKQYKKWVELPITFPNLDYSECC1	1020	Db	1 MSNGYSTDENFRYLISCFARVKMVIQVEPVLDYLTLPAEVKEIQTORTVATSGNMQAVE 60		
QY	1021	FSDED 1025	QY	421 LLNLENEDAGVQLSDFSFIIIDECHHTNKEAVNNIMRHYLMOQLKNRLLKVNKPV1P 480			
Db	1021	FSDED 1025	Db	421 LLNLENEDAGVQLSDFSFIIIDECHHTNKEAVNNIMRHYLMOQLKNRLLKVNKPV1P 480			
QY	541	1ADATREDPFKEKLLIEATMTRIQTQYQCMSPMSDFGTQTOPYQWAIQMEKKAQKGNRKERV	600	QY	481 LPQILGLTASPVGGGATKQAKAEEHILKLCANLDAFTIKTVKENIDQLKNQIQEPCKKFA 540		
Db	541	LPQILGLTASPVGGGATKQAKAEEHILKLCANLDAFTIKTVKENIDQLKNQIQEPCKKFA 540	QY	481 LPQILGLTASPVGGGATKQAKAEEHILKLCANLDAFTIKTVKENIDQLKNQIQEPCKKFA 540			

Qy	1	MSNGYSTDENFRYLISCFRAVVKMYIQLVEPVLDYLTELPAEVKEQIORTVATSGNMQAVE	60
Db	1	MSIVCSAEDSFRNLNLLFPRLKMVYIQLVEPVLDHLSAETKEQILKKNTCGNTSAAE	60
Qy	61	LLLSTLEKGWVHILGWTRFVEALRRTGSPSPLAARYMNPMLTPELSSPSENADDEYQLQLNL	120
Db	61	LLLSTLEKGWVHILGWTRFVEALRRTGSPSPLAARYMNPMLTPELSSPSENADDEYQLQLNL	120
Qy	61	LLLSTLEKGWVHILGWTRFVEALRRTGSPSPLAARYMNPMLTPELSSPSENADDEYQLQLNL	120
Db	61	LLLSTLEKGWVHILGWTRFVEALRRTGSPSPLAARYMNPMLTPELSSPSENADDEYQLQLNL	120
Qy	121	LQPTLVDKLLVRDVLDFKCMEEBLLTIEDNRNIAAENNGNEGVRELLKRVOKENWFS	180
Db	121	LQPTLVDKLLVRDVLDFKCMEEBLLTIEDNRNIAAENNGNEGVRELLKRVOKENWFS	180
Qy	121	LQPTLVDKLLVRDVLDFKCMEEBLLTIEDNRNIAAENNGNEGVRELLKRVOKENWFS	180
Db	121	LQPTLVDKLLVRDVLDFKCMEEBLLTIEDNRNIAAENNGNEGVRELLKRVOKENWFS	180
Qy	181	FLNVLROTGNELVQELTGSDCSESNAEIENLSQVDGPQVEEQLLSTTVQCPNLEKEWGM	240
Db	181	FLDVLROTGNELVQELTGSDCSESNAEIENLSQVDGPQVEEQLLSTTVQCPNLEKEWGM	240
Qy	181	FLNVLROTGNELVQELTGSDCSESNAEIENLSQVDGPQVEEQLLSTTVQCPNLEKEWGM	240
Db	181	FLDVLROTGNELVQELTGSDCSESNAEIENLSQVDGPQVEEQLLSTTVQCPNLEKEWGM	240
Qy	241	ENNSSESSPADSSVVSESDTSLAEGSVSCLDESLGHNSNMGSGTGSDSDEENV-AAR	299
Db	241	DDILPEASCTDSSVTTESDTSLAEGSVSCLDESLGHNSNMGSGTGSDSDEENV-AAR	299
Qy	241	ENNSSESSPADSSVVSESDTSLAEGSVSCLDESLGHNSNMGSGTGSDSDEENV-AAR	299
Db	241	DDILPEASCTDSSVTTESDTSLAEGSVSCLDESLGHNSNMGSGTGSDSDEENV-AAR	299
Qy	300	ASPEPELQLRPYQMEVAQPALEGKNIICLPTGSKGKTRAVYIAKDHLDDKQKASEPPGKV	359
Db	301	VSPEPELQLRPYQMEVAQPALEGKNIICLPTGSKGKTRAVYIAKDHLDDKQKASEPPGKV	359
Qy	300	ASPEPELQLRPYQMEVAQPALEGKNIICLPTGSKGKTRAVYIAKDHLDDKQKASEPPGKV	359
Db	301	VSPEPELQLRPYQMEVAQPALEGKNIICLPTGSKGKTRAVYIAKDHLDDKQKASEPPGKV	359
Qy	360	IVLVNVKVLVEQLFRKEFQFLKKWYRVIGLSGDTQLKLSFPEVVKSCDIIISTAQILEN	419
Db	361	IVLVNVKVLVEQLFRKEFQFLKKWYRVIGLSGDTQLKLSFPEVVKSCDIIISTAQILEN	419
Qy	360	IVLVNVKVLVEQLFRKEFQFLKKWYRVIGLSGDTQLKLSFPEVVKSCDIIISTAQILEN	419
Db	361	IVLVNVKVLVEQLFRKEFQFLKKWYRVIGLSGDTQLKLSFPEVVKSCDIIISTAQILEN	419
Qy	420	SLLNLENGEDAGYQLSDFSLIIDECHHTNEAVYNNIMRHYLMOKLKNRLLKENKCPV	420
Db	421	SLLNLESGDDGVQLSDDFSLIIDECHHTNEAVYNNIMRHYLMOKLKNRLLKENKCPV	420
Qy	420	SLLNLENGEDAGYQLSDFSLIIDECHHTNEAVYNNIMRHYLMOKLKNRLLKENKCPV	420
Db	421	SLLNLESGDDGVQLSDDFSLIIDECHHTNEAVYNNIMRHYLMOKLKNRLLKENKCPV	420
Qy	480	PLPQILGLTASPQVGGGATKQAKAEEHILKLCANLDAFTIKTVKENLDQLKNOIQEPCKKF	539
Db	481	PLPQILGLTASPQVGAAKKQSEAEKHILNICANLDAFTIKTVKENLDQLKNOIQEPCKKF	540
Qy	480	PLPQILGLTASPQVGGGATKQAKAEEHILKLCANLDAFTIKTVKENLDQLKNOIQEPCKKF	539
Db	481	PLPQILGLTASPQVGAAKKQSEAEKHILNICANLDAFTIKTVKENLDQLKNOIQEPCKKF	540
Qy	540	AIADATREDPFKEKLLIMTRIOTYQMSPMSPMSDFGTOPYEQWAIQMEKKAAKKGNNKERV	599
Db	541	VIADDTRENPFKEKLLIMASIQTYCQKSPMSDFGTQHYEQWAIQMEKKAAKDGNRKDVR	600
Qy	600	CAEHLRKYNEALQINDTIRMDAYTHLETIFYNEEKDKKFAVIEDDSDEGGDDEYCDGDED	659
Db	601	CAEHLRKYNEALQINDTIRMDAYSHLETFTYDKEKKFAVL-NDSDKSDDEASSCNDQI	659
Qy	660	EDDLKPLKLDLDETRFLMTLFENNKMLKRLAENPEYENKEKLTKLRTNTIMEQYTRTEESA	719
Db	660	KGDVKSKSLKLDLDEFLMNLFFDNKKMLKRLAENPKYENKEKLTKLRTNTIMEQYTRTEESA	719
Qy	720	RGIIFTKTRQSYALSQWITENEKFAEVGVKAHHLIGAGHSSEFKPMTOEQKEVTSKFR	779
Db	720	RGIIFTKTRQSYALSQWITENEKFAEVGVKAHHLIGAGHSSEFKPMTOEQKEVTSKFR	779
Qy	780	TGKINLLIATTVAAEGLDIKECNIVRYGLVITNEIAMPQARGRARADESTYVLTVAHSGSG	839
Db	780	TGEINLLIATTVAAEGLDIKECNIVRYGLVITNEIAMPQARGRARADESTYVLTSSGSG	839
Qy	840	VIEHETVNDFREKMYKAITHCVQNMQPEYAHKILLEQMTQEQKEVTSKFR	959
Db	840	VTERETVNDFREKMYKAITHCVQNMQPEYAHKILLEQMTQEQKEVTSKFR	959
Qy	900	PSLITFLCKNCVSPLACSGDIIHVIKEKMHVNMTPFEFKGLYIVRENKALQKCFADYQTNGE	959
Db	900	PSLITFLCKNCVSPLACSGDIIHVIKEKMHVNMTPFEFKGLYIVRENKALQKCFADYQTNGE	959
Qy	960	LIICKCGQAWGTMVHKGDLPLCKNCSMMLVCSGENIHVIKEKMHVNMTPFEFKGLYIVRENKALQKCFADYQTNGE	1019
Db	960	LIICKCGQAWGTMVHKGDLPLCKNCSMMLVCSGENIHVIKEKMHVNMTPFEFKGLYIVRENKALQKCFADYQTNGE	1019
Qy	1020	LFSDED 1025	
Db	1020	LFSDED 1025	
Qy	1020	LFSDED 1025	
Db	1020	LFSDED 1025	
Qy	1020	Query Match 79.1%; Score 4199; DB 11; Length 1025;	
Qy	1020	Best Local Similarity 79.7%; Pred. No. 3.6e-226;	
Qy	1020	Matches 818; Conservative 75; Mismatches 131; Indels 2; Gaps 2;	

RESULT	4	Db	481 PLPQILGLTASPGVGAAKQSEAKHILNICALDAFTIKTVKENLQLKHQIKEPCKKF	540
Q8K5C7	PRELIMINARY;	PRT;	1025 AA.	
ID	Q8K5C7;			
AC	Q8K5C7;			
DT	01-OCT-2002 (TREMBLrel. 22, Created)			
DT	01-OCT-2002 (TREMBLrel. 22, Last sequence update)			
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)			
DE	Melanoma differentiation associated gene 5-like protein.			
GN	MDA5 OR 913009C22RIK.			
OS	Mus musculus (Mouse)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6J;			
RA	Kang, D. -C.; Fisher, P.B.;			
RT	"Murine homolog of melanoma differentiation associated gene-5 - a DEXH group RNA dependent ATPase.";			
RT	Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.			
RL	EMBL; AF374384; FAM21359.1; -.			
DR	MGI:1918836; Mda5.			
DR	GO; GO:0005524; F:ATP binding; IEA.			
DR	GO; GO:0008026; F:ATP dependent helicase activity; IEA.			
DR	GO; GO:0016787; F:hydrolease activity; IEA.			
DR	GO; GO:0003676; F:nucleic acid binding; IEA.			
DR	DR InterPro; IPR001410; DEAD.			
DR	DR InterPro; IPR001650; Helicase_C.			
DR	PFam; PF00270; DEAD; 1.			
DR	PFam; PF00271; helicase_C; 1.			
DR	SMART; SM000487; DEXDC; 1.			
DR	SMART; SM000490; HELICC; 1.			
KW	ATP-binding; Helicase; Hydrolase.			
SEQUENCE	1025 AA; 115941 MW; 8338A8510C5ED3AB CRC64;			
SQ				
Query Match	79.0%;	Score 4194;	DB 11;	Length 1025;
Best Local Similarity	79.6%;	Pred. No. 6.9e-226;		
Matches	817;	Conservative 75;	Mismatches 132;	Indels 2;
Matches	817;		Gaps 2;	
Query	1 MSNGYSTDENFRLISCFRARVAMYIQVEPVYDYLTLPEAIVEKEQIORTVATSGNMQAVE	60		RESULT 5
Db	1 MSIVCSAEDDSFRNLILFFRPRLMMYIQVEPVLDLIFLSAETKEQILKKINTCGNTSAAE	60		Q8BZ01
QY	1 LLLSTLEKGWWHLGWTREFVEALLRTGSPLAARYMNPELTDLPSFSENAHDEYLQLLNL	120		PRELIMINARY;
Db	61 LLLSTLEKGWWHLGWTREFVEALLRTGSPLAARYMNPELTDLPSFSENAHDEYLQLLNL	120		PRT;
QY	61 LLLSTLEKGWWHLGWTREFVEALLRTGSPLAARYMNPELTDLPSFSENAHDEYLQLLNL	120		693 AA.
Db	121 LQPTLVDKLLYVRDVLDKCMEEELLTIEDRNRRIAAENNGNEGSVRELLKRIQKENWFA	180		
QY	121 LQPTLVDKLLYVRDVLDKCMEEELLTIEDRNRRIAAENNGNEGSVRELLKRIQKENWFA	180		Q8BZ01
Db	121 LQPTLVDKLLYVRDVLDKCMEEELLTIEDRNRRIAAENNGNEGSVRELLKRIQKENWFA	180		ID
QY	181 FFLNVLRQTGNNLVQELTRDNRRIAAENNGNEGSVRELLKRIQKENWFA	180		Q8BZ01;
Db	181 FFLNVLRQTGNNLVQELTRDNRRIAAENNGNEGSVRELLKRIQKENWFA	180		AC
QY	181 FFLNVLRQTGNNLVQELTRDNRRIAAENNGNEGSVRELLKRIQKENWFA	180		01-MAR-2003 (TREMBLrel. 23, Created)
Db	181 FFLNVLRQTGNNLVQELTRDNRRIAAENNGNEGSVRELLKRIQKENWFA	180		DT
QY	241 ENNSSESSFADSSVVSESDTSLAEGSVSCLDESIGHNSNMGSDSDEENV-AAR	299		01-MAR-2003 (TREMBLrel. 23, Last sequence update)
Db	241 DDLIPEASCTDSSVTTESDTSLAEGSVSCLDESIGHNSNMGRDSGTMGSDSDESVIQTKR	300		DT
QY	241 ENNSSESSFADSSVVSESDTSLAEGSVSCLDESIGHNSNMGSDSDEENV-AAR	299		01-OCT-2003 (TREMBLrel. 25, Last annotation update)
Db	241 ENNSSESSFADSSVVSESDTSLAEGSVSCLDESIGHNSNMGRDSGTMGSDSDESVIQTKR	300		DE
QY	300 ASPEPELQLRQYQMEVAQPALEGKNIICLPTGSKTRAVAYTAKDHLDDKKKQASESGKV	359		Similar to melanoma differentiation associated protein-5 (Fragment).
Db	301 VSPEPELQLRQYQMEVAQPALEGKNIICLPTGSKTRAVAYTAKDHLDDKKKQASESGKV	360		DE
QY	360 IVLVNKVLLVEOLFRKEFQPFLLKXRVIGLSGDQTQLKISFPEVVKSCDIIISTAQILEN	419		GN
Db	361 IVLVNKVLLVEOLFRKEFQPFLLKXRVIGLSGDQTQLKISFPEVVKSCDIIISTAQILEN	420		MDA5 OR 9130009C22RIK.
QY	420 SLLNLENGEDAGYQLSDFSLIIIDECIHTNKEAVYNNIMRHYLMQKLKNRLLKQENKPKV	479		OS
Db	421 SLLNLENGEDAGYQLSDFSLIIIDECIHTNKEAVYNNIMRHYLMQKLKNRLLKQENKPKV	480		Eukaryota; Metazoa; Craniata; Vertebrata; Euteleostomi;
QY	480 PLPQILGLTASPGVGAAKQSEAKHILNICALDAFTIKTVKENLQLKHQIKEPCKKF	539		OC
Db	481 PLPQILGLTASPGVGAAKQSEAKHILNICALDAFTIKTVKENLQLKHQIKEPCKKF	540		Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
QY	540 AIADATREDPFREKLLLEIMTRIQTQYBQWAIQMEKKAARKKGNRKERV	599		NCBI_TaxID=10090;
Db	541 VIADDTRENPFKEKLLIMASIQTQYBQWAIQMEKKAARKKGNRKDRV	600		[1]
QY	600 CAEHLRKYNEALQINDTIRMIDAYTLETFYNEEKDKKKPAVIEDDSDEGGDDED	659		
Db	601 CAEHLRKYNEALQINDTIRMIDAYSHLEAFYTDKEKKPAVFL-NDSDKSDDEASSCNDQL	659		
QY	660 EDDLKPKPLKLDDETDRFLMLTFFENNKMKLKRLAENPEYENENKLTKLRTNTIMEQYTRTEESA	719		
Db	660 KGDKVKSLKLKDDETDEFMLNLFFDNKMKMLKKLAAENPKYENENKLTKLRTNTILEQFTTRSEES	719		
QY	720 RGIIIFTKTRQSAVALSQWITENEKFAEVGYKAHLIGAGHSSEFKPMTQNEQKEVISKFR	779		
Db	720 RGIIIFTKTRQSTYALSQWIMENAKFAEVGVKAHLIGAGHSSEVKPMTQTEQKEVISKFR	779		
QY	780 TGKINLLIATTVAEEEGLDIKECNVIRYGLVTNEIAMQARGRADAESTYVLYAHSGSG	839		
Db	780 TGEINLLIATTVAEEGLDIKECNVIRYGLVTNEIAMQARGRADAESTYVLYVTSSGSG	839		
QY	840 VIEHETVNDFREKMTYKAHCVQNMKPEEYAHKILELQMSIMEKOMKTKRNLAKHYQNN	899		
Db	840 VTERETVNDFREKMTYKAIRVQNMKPEEYAHKILELQVQSTILEKKMKVKRSTAKQYNDN	899		
QY	900 PSLITFLCKNCSVVLACSGEDIHVIEKMEHVNMTPEFKELYIVRENKALQKFCADYQINGE	959		
Db	900 PSLITFLCKNCSMLYCSEGENTHVIEKMHMHTVNMTPFEGKLYIVRENKALQKFCADYQTNGE	959		
QY	960 IICKCGQAWGTMAMVHKGLDLPCLKIRNFVUNFKNNNSPKQYKKWVLPIRFPDLDYSEYCC	1019		
Db	960 IICKCGQAWGTMAMVHKGLDLPCLKIRNFVUNFKNNNSPKQYKKWVLPIRFPDLDYSEYCC	1019		
QY	1020 LFSDED 1025			
Db	1020 LYSDED 1025			
QY	1 MSNGYSTDENFRLISCFRARVAMYIQVEPVYDYLTLPEAIVEKEQIORTVATSGNMQAVE	60		RESULT 5
Db	1 MSIVCSAEDDSFRNLILFFRPRLMMYIQVEPVLDLIFLSAETKEQILKKINTCGNTSAAE	60		Q8BZ01
QY	61 LLLSTLEKGWWHLGWTREFVEALLRTGSPLAARYMNPELTDLPSFSENAHDEYLQLLNL	120		PRELIMINARY;
Db	61 LLLSTLEKGWWHLGWTREFVEALLRTGSPLAARYMNPELTDLPSFSENAHDEYLQLLNL	120		PRT;
QY	121 LQPTLVDKLLYVRDVLDKCMEEELLTIEDRNRRIAAENNGNEGSVRELLKRIQKENWFA	180		693 AA.
Db	121 LQPTLVDKLLYVRDVLDKCMEEELLTIEDRNRRIAAENNGNEGSVRELLKRIQKENWFA	180		
QY	181 FFLNVLRQTGNNLVQELTRDNRRIAAENNGNEGSVRELLKRIQKENWFA	180		
Db	181 FFLNVLRQTGNNLVQELTRDNRRIAAENNGNEGSVRELLKRIQKENWFA	180		
QY	241 ENNSSESSFADSSVVSESDTSLAEGSVSCLDESIGHNSNMGSDSDEENV-AAR	299		
Db	241 DDLIPEASCTDSSVTTESDTSLAEGSVSCLDESIGHNSNMGRDSGTMGSDSDESVIQTKR	300		
QY	300 ASPEPELQLRQYQMEVAQPALEGKNIICLPTGSKTRAVAYTAKDHLDDKKKQASESGKV	359		
Db	301 VSPEPELQLRQYQMEVAQPALEGKNIICLPTGSKTRAVAYTAKDHLDDKKKQASESGKV	360		
QY	360 IVLVNKVLLVEOLFRKEFQPFLLKXRVIGLSGDQTQLKISFPEVVKSCDIIISTAQILEN	419		
Db	361 IVLVNKVLLVEOLFRKEFQPFLLKXRVIGLSGDQTQLKISFPEVVKSCDIIISTAQILEN	420		
QY	420 SLLNLENGEDAGYQLSDFSLIIIDECIHTNKEAVYNNIMRHYLMQKLKNRLLKQENKPKV	479		
Db	421 SLLNLENGEDAGYQLSDFSLIIIDECIHTNKEAVYNNIMRHYLMQKLKNRLLKQENKPKV	480		
QY	480 PLPQILGLTASPGVGAAKQSEAKHILNICALDAFTIKTVKENLQLKHQIKEPCKKF	539		
Db	481 PLPQILGLTASPGVGAAKQSEAKHILNICALDAFTIKTVKENLQLKHQIKEPCKKF	540		
QY	540 AIADATREDPFREKLLLEIMTRIQTQYBQWAIQMEKKAARKKGNRKERV	599		
Db	541 VIADDTRENPFKEKLLIMASIQTQYBQWAIQMEKKAARKKGNRKDRV	600		
QY	600 CAEHLRKYNEALQINDTIRMIDAYTLETFYNEEKDKKKPAVIEDDSDEGGDDED	659		
Db	601 CAEHLRKYNEALQINDTIRMIDAYSHLEAFYTDKEKKPAVFL-NDSDKSDDEASSCNDQL	659		
QY	660 EDDLKPKPLKLDDETDRFLMLTFFENNKMKLKRLAENPEYENENKLTKLRTNTIMEQYTRTEESA	719		
Db	660 KGDKVKSLKLKDDETDEFMLNLFFDNKMKMLKKLAAENPKYENENKLTKLRTNTILEQFTTRSEES	719		
QY	720 RGIIIFTKTRQSAVALSQWITENEKFAEVGYKAHLIGAGHSSEFKPMTQNEQKEVISKFR	779		
Db	720 RGIIIFTKTRQSTYALSQWIMENAKFAEVGVKAHLIGAGHSSEVKPMTQTEQKEVISKFR	779		
QY	780 TGKINLLIATTVAEEEGLDIKECNVIRYGLVTNEIAMQARGRADAESTYVLYAHSGSG	839		
Db	780 TGEINLLIATTVAEEGLDIKECNVIRYGLVTNEIAMQARGRADAESTYVLYVTSSGSG	839		
QY	840 VIEHETVNDFREKMTYKAHCVQNMKPEEYAHKILELQMSIMEKOMKTKRNLAKHYQNN	899		
Db	840 VTERETVNDFREKMTYKAIRVQNMKPEEYAHKILELQVQSTILEKKMKVKRSTAKQYNDN	899		
QY	900 PSLITFLCKNCSVVLACSGEDIHVIEKMEHVNMTPEFKELYIVRENKALQKFCADYQINGE	959		
Db	900 PSLITFLCKNCSMLYCSEGENTHVIEKMHMHTVNMTPFEGKLYIVRENKALQKFCADYQTNGE	959		
QY	960 IICKCGQAWGTMAMVHKGLDLPCLKIRNFVUNFKNNNSPKQYKKWVLPIRFPDLDYSEYCC	1019		
Db	960 IICKCGQAWGTMAMVHKGLDLPCLKIRNFVUNFKNNNSPKQYKKWVLPIRFPDLDYSEYCC	1019		
QY	1020 LFSDED 1025			
Db	1020 LYSDED 1025			
QY	1 MSNGYSTDENFRLISCFRARVAMYIQVEPVYDYLTLPEAIVEKEQIORTVATSGNMQAVE	60		RESULT 5
Db	1 MSIVCSAEDDSFRNLILFFRPRLMMYIQVEPVLDLIFLSAETKEQILKKINTCGNTSAAE	60		Q8BZ01
QY	61 LLLSTLEKGWWHLGWTREFVEALLRTGSPLAARYMNPELTDLPSFSENAHDEYLQLLNL	120		PRELIMINARY;
Db	61 LLLSTLEKGWWHLGWTREFVEALLRTGSPLAARYMNPELTDLPSFSENAHDEYLQLLNL	120		PRT;
QY	121 LQPTLVDKLLYVRDVLDKCMEEELLTIEDRNRRIAAENNGNEGSVRELLKRIQKENWFA	180		693 AA.
Db	121 LQPTLVDKLLYVRDVLDKCMEEELLTIEDRNRRIAAENNGNEGSVRELLKRIQKENWFA	180		
QY	181 FFLNVLRQTGNNLVQELTRDNRRIAAENNGNEGSVRELL			

Best Local Similarity	75.6%	Pred. No.	3.8e-141;
Matches	525;	Conservative	60;
Indels	2;	Gaps	2;
RT	60,770	full-length	CDNA8. ";
RL	Nature	420:563-573	(2002).
DR	EMBL;	AK040519;	BAC30614.1;
DR	MGD;	MGI:1918836;	Mda5.
DR	GO;	GO:0005524;	F-ATP binding;
DR	GO;	GO:0008026;	F-ATP dependent helicase activity;
DR	InterPro;	IPR001410;	DEAD.
DR	SMART;	SM00487;	DEXD.
FT	NON TER	683	683
SQ	SEQUENCE	683 AA;	76572 MW;
			012F0414B28251C3 CRC64;
Query	1	MSNGYSTDENFRYLISCFRARVKMYIQVEPVYDYLTFPAEVKEQIQRTVATSGNMQAVE	60
Db	1	MSIVCSAEDSFRLILFERRPRLKMYIQVEPVYDHLIFLSAETKEQILRKINTCGNTSAAE	60
Query	61	LLLSTLEKGWHLGWTRFEALRRRTGSPLAARYMNPELTDLPSFSFENAHDDEYLQLLNL	120
Db	61	LLLSTLEKGWHLGWTSSETAHDDECLHLLTL	120
Query	121	LQPTLVDKLLVRDVLDLKDCKMEEELLTTEDRNRIAAEENNNGSVRELLKRIVKOKENWFA	180
Db	121	LQPTLVDKLLINDVLDLTCFEKGLLTVEDRNRTSAAGNSGNESGVRELLRRIVKOKENWFST	180
Query	181	FLNVLRQTGNNELVQELTGSDCSESNAE1ENLSQLVDGPQVEBQLLTTVOPNLEKEVWGM	240
Db	181	FLDVLRQTGNDALFQBLTGGGCPEDNTDLANSSSHRGPAANECLIPAVDESSLETEAWN	240
Query	241	ENNSSESSFADSSVVSESDTSLAEGSVSCLDES LGHNSNMGSDSGTMGSDSDEENV-AAR	299
Db	241	DDILPEASCTDSSVTTESTDTSLAEGSVSFCDESLGHNSNMGRDSGTMGSDSDESVIQTKR	300
Query	300	ASPEPELQLRPyQMEVAQPALEGKNNIIICLPTGSGKTRVAVYIAKDHDLKKKASEPGK	359
Db	301	VSPEPELQLRPyQMEVAQPALDGKNNIIICLPTGSGKTRVAVYITKDHDLKKKQASESGK	360
Query	360	IVLVNKVLLVEQLFRKEFQPFLLKKWYRVLGIGLQDGTOLKISFPEVVKSCDIIISTAQILEN	419
Db	361	IVLVNKVMLAEQLFRKEFNPYLLKKWYRIGLQDGTOLKISFPEVVKSYDVIISTAQILEN	420
Query	420	SLLNLENGEDAGVQLSDFSLIIIDECHTNKEAVYNNIMRHYLMQKLKNRLLKKENKPVI	479
Db	421	SLLNLESQDDDGQVLSDFSLIIIDECHTNKEAVYNNIMRRLYKQKLRNNDLKKQNKPAI	480
Query	480	PLPQILQJLTASPGVGATKQAKAEEHLKUCANLDAFTIKTVKENLQKLQKQNKPAI	539
Db	481	PLPQILQJLTASPGVGAAKKQSEAEKHILNICANLDAFTIKTVKENLQKLQKQNKPAI	540
Query	540	AIADATREDPFEKLEIMTRIQTYCQMSPMSDFGTQPYEOWAIQMEKKAAKKGNRKERV	599
Db	541	VIADDTRENPFKEKLEIMASIQTYCQKSPMSDFGTQHYEOWAIQMEKKAAKDGNRKDRV	600
Query	600	CAEHLRKYNEALQINDTIRMIDAYTHLETFYNEBKOKKFAVIEDDSDEGGDDEYCDGDED	659
Db	601	CAEHLRKYNEALQINDTIRMIDAYSHLETFTDEKEKKFAVL-NDSDKSDEASSCNDQL	659
Query	660	EDDLKKPLKLDDETDFLMLTFENNKMVKLRAEN	693
Db	660	KGDVKKSILKLDDETDFLMNLFFDNKKMLKKLAAEN	693
RESULT	6		
Q8BYC9		PRELIMINARY;	PRT;
Q8BYC9		683 AA.	
AC			
DT	01-MAR-2003	(TREMBLrel. 23, Created)	
DT	01-MAR-2003	(TREMBLrel. 23, Last sequence update)	
DT	01-OCT-2003	(TREMBLrel. 25, Last annotation update)	
DE		Similar to melanoma differentiation associated protein-5 (Fragement).	
DE		MDA5 OR 9130009C22RIK.	
GN		Mus musculus (Mouse).	
OC		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC		Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
NCBI_TaxID	10090;		
RN			RESULT 7
			Q96MX8
RP	SEQUENCE FROM N.A.		ID Q96MX8
RC	STRAIN=C57BL/6J; TISSUE=Thymus;		AC Q96MX8;
RX	MEDLINE=22354683; PubMed=12466851;		DT 01-DEC-2001 (TREMBLrel. 19, Created)
RA	The FANTOM Consortium,		DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
RA	"Analysis of the mouse transcriptome based on functional annotation of		DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
RT			DE Hypothetical protein FLJ31731.



SEQUENCE FROM N.A.	
Strausberg, R.	Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
EMBL; BC025508; AAH25508.1;	-
GO; GO:0005524; F:ATP binding; IEA.	
GO; GO:0008026; F:ATP dependent helicase activity; IEA.	
GO; GO:0016787; F:hydrolase activity; IEA.	
GO; GO:0003676; F:nucleic acid binding; IEA.	
InterPro; IPR001410; DEAD.	
InterPro; IPR001650; Helicase_C.	
PFam; PF00271; helicase_C; 1.	
SMART; SM00490; HELIC_C; 1.	
Hypothetical protein; ATP-binding; Helicase; Hydrolase.	
SEQUENCE 467 AA; 53966 MW; F3B0D976778F0442 CRC64;	
Query Match 38.8%; Score 2059.5; DB 11; Length 467;	
Best Local Similarity 84.6%; Pred. No. 4.2e-107;	
Matches 396; Conservative 28; Mismatches 43; Indels 1; Gaps 1	
MTRIQTYCQMSPMSDFGTOPYEQWAQMEKKAAKKGNRICKVCAEHLRKYNEALQINDTI	
1 MASIQTYCQKSPMSDFGTQHXYEQWAQMEKKAAKDGNRICKVCAEHLRKYNEALQINDTI	
618 RMIDAYTHILETFYNEEKDKKKPAVIEDDSDEGGDDEYCDGDEDDEDDLKKPLKLKDETDRFLM	
61 RMIDAHSHLJETFYTDKEKKPAVL-NDSDESSDEASSCNDQLKGDVKSLSLKLDDETDEFLM	
678 TLFENNKKMLKRLAENPEYENEEKLTKLRNTIMEQYTRTEESARGIIFTKTRQSAVYALSQW	
120 NLFFFDNKQQLKLAENPKYENEEKLKLIKLRNTILEQFTTRSEESSRGIFTKTRQSTYALSQW	
738 ITENEKPAEVGVKAHHLIGAHSSEFKPMTQNEQKEVISKFRTGKINLLIATTVAAEGLD	
180 IMENAKFAEVGVKAHHLIGAHSSEVKPMTQTEQKEVISKFRTGIEINLLIATTVAAEGLD	
798 IKECNIVIRYGLVTNEIAMVQARGRARADESTYVLVAHSGGVIEHETVNDFREKMYKA	
240 IKECNIVIRYGLVTNEIAMVQARGRARADESTYVLVTSSGGVTTEREVNDFREKMYKA	
858 IHCVQNMKPEEYAHKILLEQMSITMEKKMKTKRNTIAKHYKNNPSSLITFLCKNCSVLACSG	
300 INRVQNMKPEEYAHKILELQVOSILEKQMKVKRSIAKQYNDNPSLITLLCKNCSMLVCSG	
918 EDIHVIEKMHVNMTPEFKELYITRENKAHQKKCADYQINGEIIICKCQOAWTMMVHKGJ	
360 ENHVIEKMHVNMTPEFKGLYITRENKAHQKKFADYQINGEIIICKCQOAWTMMVHKGJ	
978 DLPCLKIRNFVYVFKNNSTKKQYKKWELPITFPNLDYSECCCLFSDED 1025	
420 DLPCLKIRNFVYVFKNNSPKQYKKWELPIRFPDLDYSECYCLYSDED 467	
SULT 10	
9J78 999JB7 PRELIMINARY; PRT; 678 AA.	
Q99JB7; 01-JUN-2001 (TREMBLrel. 17, Created)	
01-JUN-2001 (TREMBLrel. 17, Last sequence update)	
01-OCT-2003 (TREMBLrel. 25, Last annotation update)	
Hypothetical protein (DNA segment, Chr 11, Lothar Hennighausen 2, expressed).	
DLLGP2E OR DLLGP2.	
Mus musculus (Mouse)	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
NCBI_TaxID=10090;	
[1]	
SEQUENCE FROM N.A.	
PubMed=11161808;	





AC	095786;	Db	353	N--NLKKGTTIP--SLSIFTLMI	DECHNTSKQHFPYNMIMENYLQDKLGSS	-----	399
DT	01-MAY-1999 (TREMBLrel. 10, Created)	Qy	479	IPLQIQLGLTASPVGATKQAKAEHILKLCANLDAFTIKTVKENILDQLKNQIQEPCKK	538		
DT	01-MAY-1999 (TREMBLrel. 10, Last sequence update)	Db	400	GPLPQVIGLTASVGVDAXNTDEALDYICKLQASLDAVIAVCHNLLEEQVYKPKQKF	459		
DE	RNA helicase.						
GN	RIG-I.						
OS	Homo sapiens (Human).						
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.						
OX							
NCBI_TAXID	9606;						
[1]	SEQUENCE FROM N.A.						
RA	Sun Y.W.;						
RA	"RIG-I", a human homolog gene of RNA helicase, is induced by retinoic acid during the differentiation of acute promyelocytic leukemia cell.";						
RT	Thesis (1997), Shanghai Institute of Hematology, Rui-Jin Hospital, Shanghai Second Medical University.						
RN	[12]						
RP	SEQUENCE FROM N.A.						
RA	Yi-Wu S.;						
RL	Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.						
DR	EMBL; AF038963; ADD19826.1; -.						
DR	GO; GO:0005524; F:ATP binding; IEA.						
DR	GO; GO:0008026; F:ATP dependent helicase activity; IEA.						
DR	GO; GO:0016787; F:hydrolase activity; IEA.						
DR	GO; GO:0003676; F:nucleic acid binding; IEA.						
DR	GO; GO:0000166; F:nucleotide binding; IEA.						
DR	InterPro; IPR003593; AAA; ATPase.						
DR	InterPro; IPR001410; DEAD.						
DR	InterPro; IPR001650; Helicase_C.						
DR	Pfam; PF00270; DEAD; 1.						
DR	Pfam; PF00271; helicase_C; 1.						
DR	SMART; SM00382; AAA; 1.						
DR	SMART; SM00487; DEXDC; 1.						
DR	SMART; SM00490; HELICC; 1.						
KW	ATP-binding; Helicase; Hydrolase.						
SQ	SEQUENCE 925 AA; 106613 MW; 5922B4F3DD0F00D4 CRC64;						
Query Match	20.5%; Score 1086.5; DB 4; Length 925;						
Best Local Similarity	31.2%; Pred. No. 2.4e-52;						
Matches	325; Conservative 166; Mismatches 364; Indels 185; Gaps 35;						
Qy	12 RYLSICFRARVKMYIQVEPVLDYLT--FLPAEVKEQIQRSTVATSGNMQAVERLLSTLEKG	69	RESULT 15				
Db	6 RRSLSQAFQDYIRKTLDPYIISYMAPWFREEV-QYIQAERKNNKGPMEAATLFLKFLL-63	Q86X56	PRELIMINARY;		PRT;	221 AA.	
Qy	70 VWHLGWTREFVEALLRTGSPSPLAARYMNPELTDLPSPSFENAH--DEYLQLNLLQPTLVD	127	AC	Q86X56;			
Db	74 LQEEGWFGRGFLDAHDAG-YSGLYEALESWD---FKKIEKLEYYRLLRQLQPEFKT	116	DT	01-JUN-2003 (TREMBLrel. 24, Created)			
Qy	128 KLLVRDVLDKCMEEELLTIEDRNRRIAAEENNNGNESGVRELLKRIVQ--KENWFSAFNLVL	185	DT	01-JUN-2003 (TREMBLrel. 24, Last sequence update)			
Db	117 RIPTDIISD-LSECECLINQCEEEILOQICSTKGMAGAEKLVECLLRSKDKENWPKTL---	171	DT	01-JUN-2003 (TREMBLrel. 24, Last annotation update)			
Qy	186 RQTGNNEVQELTGSDCESNAEIEENLSQVDGPQVEEQLLSTTVQPNLEKEWIGMEN--	242	DE	Similar to melanoma differentiaton associated protein-5.			
Db	172 -----KLALEKERNKFSSELWIVKGKIDVETEDLDRKMETSDIQL-----YS 239	219	OS	Homo sapiens (Human).			
Qy	243 -NSSESSFADSSVVSESNTSLAEGSVSCLDESLGHNSNMGSDDSGTMGSDEENVAARAS	301	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Db	220 QNLSENSCPPSEV--SDTNL-----YS 239	219	OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
Qy	302 PEPQLQRPYOMEVAQPALEGKNIICLPTGSKTRVAVYIAKDHLDKKKKASEPGKVI	361	OX	NCBI_TaxID=9606;			
Db	240 P---FKERNYQOLELLALPAMKGKNTIICAPTGGKTFVSLICEHHL-KKFPQGQKGKVV	295	RN	[1]			
Qy	362 LVNKVLLVEQ--LFRKEFQPELKKWYRVLGLSGDTQLKISFPEVKSCDIIISTAQILE	418	RA	Strausberg R.;			
Db	296 FANQIPVYEQOKSVFSKYF--RHGYRTGSGATAENVPEQIVENNDIIILTPQILV	352	RI	Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.			
Qy	419 NSLLNLLENGDAGVQLSDFSLIIDECHHTNKEAVYNTMRHYLMQKLKNRUKKENKPV	478	DR	EMBL; BC04620B; AAH46208.1;			
Db		295	SQ	SEQUENCE 221 AA; 25129 MW; 3C7139ECAC564BB5 CRC64;			
Qy		295	Query Match	20.2%; Score 1071; DB 4; Length 221;			
Db		295	Best Local Similarity	97.2%; Pred. No. 2.6e-52;			
Matches	210; Conservative 2; Mismatches 4; Indels 0; Gaps 0;						
Qy	1 MSNGYSTDENFRYLISCFRARVKMYIQVEPVLDLTFLPAEVKQIORTVATSGNMQAVE	60					
Db	1 MSNGYSTDENFRYLISCFRARVKMYIQVEPVLDLTFLPAEVKQIORTVATSGNMQAVE	60					
Qy	61 LLLSTLEKGWHLIGWTREFVEALRRTGSPSLEARYMNPELTDLPSPSFENAHDEYIQLJNL	120					

Db	61	LILSTLEKGWVHLGWTREFVEALRRRTGSPLAARYMNPELTDLPSPEFENAHDEYLQLLNL	120
Qy	121	LQPTLVDKLLVRDVLDKCMEEELLTTIEDRNRIAAENNGNESGVRELLKRTIVKENWFS	180
Db	121	LQPTLVDKLLVRDVLDKCMEEELLTTIEDRNRIAAENNGNESGVRELLKRTIVKENWFS	180
Qy	181	FLNVLRQIGNNELVQELTGSDCSESNAEIEIENLSQVD	216
Db	181	FLNVLRQIGNNELVQELTGSDCSESNAEIEIENLSQVD	216

Search completed: June 2, 2004, 19:11:47  
Job time : 53 secs